


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: NASA M&S Standard			Page #: i

NASA Standard for Models and Simulations (M&S): Development Process and Rationale

Response to Action Item #4 from the *A Renewed Commitment to Excellence: An Assessment of the NASA Agency-wide Applicability of the Columbia Accident Investigation Board Report*

November 20, 2008

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: ii

Report Approval and Revision History

NOTE: This document was approved at the November 20, 2008, NRB. This document was submitted to the NESC Director on November 21, 2008, for configuration control.

Approved:	Original signature on file. RR	11-24-08
	NESC Director	Date

Revision	Description of Revision	Author	Effective Date
1.0	Initial Release	Thomas A. Zang, Development Team Lead & Topic Working Group Chair	11-20-08



	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: iii

Table of Contents

Volume I: Technical Assessment Report

1.0	Notification and Authorization	1
2.0	Signature Page.....	2
3.0	Team List	2
3.1	Acknowledgements	4
4.0	Executive Summary	5
5.0	Assessment Plan	6
6.0	Problem Description and Scope.....	8
6.1	Columbia Accident Investigation Board	8
6.2	A Renewed Commitment to Excellence: An Assessment of the NASA Agency-wide Applicability of the Columbia Accident Investigation Board Report	9
6.3	NASA Chief Engineer Guidance	10
7.0	Major Activities.....	11
7.1	Timeline of M&S Standard Development.....	11
7.2	Phase 1 Activities	13
7.2.1	Review of Existing M&S Guidance and Standards	13
7.2.2	Space Shuttle Orbiter RCC Impact Damage Threshold Assessment Team Pilot Study	15
7.2.3	M&S Scales/Maturity Matrices	16
7.2.4	First Three Versions of the Standard	17
7.3	The Interim M&S Standard.....	20
7.3.1	Background on Interim M&S Standards.....	20
7.3.2	Version 3 Review.....	21
7.3.3	Rollout of the Interim M&S Standard	23
7.3.4	Pilot Studies on the Interim M&S Standard	23
7.4	Phase 2 Scale Workshops.....	26
7.4.1	Scale Workshop 1	27
7.4.2	Scale Workshop 2	29
7.4.3	Scale Workshop 3	32
7.4.4	Scale Workshop 4	35
7.4.5	Sandia National Laboratories-NASA Workshop.....	36
7.5	Formal Agency-wide Review.....	37
7.5.1	Interim M&S Standard Review	37
7.5.2	Major Changes in the Revised M&S Standard	39
7.6	Engineering Management Board Review.....	41

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: iv

7.6.1	Major Changes in the Permanent M&S Standard	41
7.6.2	Risk Assessment Matrix	43
7.6.3	Issuance of Permanent M&S Standard	43
7.7	Outreach	44
8.0	Overview of the Permanent M&S Standard	45
8.1	Scope	46
8.2	Definitions	48
8.3	Requirements	48
8.4	Credibility Assessment Scale	49
8.5	Traceability	50
9.0	Observations, Findings and Recommendations	51
9.1	Summary of Observations and Findings	51
9.2	Development Team Recommendations	52
9.3	Topic Working Group Recommendations	53
10.0	Alternate Viewpoints	55
10.1	Alternate Viewpoint (Unmeel Mehta) on the Interim M&S Standard	55
10.2	Alternate Viewpoint (Unmeel Mehta) on the Revised M&S Standard	55
10.3	Alternate Recommendation (Unmeel Mehta)	55
11.0	Other Deliverables	56
11.1	Standard Documents	56
11.2	Pilot Scale Questionnaire	57
11.3	Decision-Maker Interviews	59
11.4	Pilot Questionnaire	59
12.0	Lessons Learned	62
13.0	Definition of Terms	62
14.0	Acronyms List	63
15.0	References	64

List of Figures

Figure 5.0-1.	Phase 1 of the M&S Standard Development	7
Figure 5.0-2.	Phase 2 of the M&S Standard Development	7
Figure 7.4-3.	M&S Credibility Assessment Scale Development	27
Figure 7.4-4.	Brainstorming on the Best Architecture	30
Figure 7.4-5.	Brainstorming on the Key Factors for Assessing Credibility	31
Figure 7.4-6.	Initial Evaluation of Potential Credibility Assessment Scale Factors	33
Figure 7.6-7.	Sample M&S Risk Assessment Matrix	43


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: v


Figure 8.0-1. Two Parts of the M&S Standard	45
Figure 11.2-1. Bar Chart of Results from the Pilot Scale Questionnaire	58
Figure 11.2-2. Mosaic Plot of Results from the Pilot Scale Questionnaire	58
Figure 11.4-1. Estimate of Cost to Achieve Level 1 on the A2 Scale	60
Figure 11.4-2. Estimate of Cost to Achieve Level 3 on the A2 Scale	61

List of Tables

Table 7.1-1. M&S Standard Phase 1 Timeline	11
Table 7.1-2. M&S Standard Phase 2 Timeline	12
Table 7.3-3. Topic Working Group Vote on the Interim M&S Standard	22
Table 7.3-4. Pilot Studies Using the Interim M&S Standard	23
Table 7.5-5. Topic Working Group Votes on the Revised M&S Standard	38
Table 11.4-1. Raw Data for Cost Estimates for Level 1	60

Volume II Appendices

Appendix A. NASA Chief Engineer Memo	67
Appendix B. Excerpts from the Space Shuttle Return to Flight Report	71
Appendix C. ARC Topic Working Group Member's Objections to the Interim M&S Standard	72
Appendix D. Consequence Definitions from NPR 8000.4	75
Appendix E. Traceability Matrix	76
Appendix F. Pilot Scale Questionnaire	99
Appendix G. Decision Maker Interview Guide	106
Appendix H. Pilot Questionnaire	108

	NASA Engineering and Safety Center Technical Report	Document RP-08-118	Version: 1.0
Title:	M&S Standard Completion		Page #: 1 of 117

Volume I: Technical Assessment Report

1.0 Notification and Authorization

In March 2005, Mr. Michael Blythe of the Office of the Chief Engineer (OCE) requested the Director, Systems Analysis and Concepts Directorate at the Langley Research Center (LaRC) to submit a plan for responding to Diaz Action #4 from the report *A Renewed Commitment to Excellence: An Assessment of the NASA Agency-wide Applicability of the Columbia Accident Investigation Board Report* [ref. 1].


“Develop a standard for the development, documentation, and operation of models and simulations.

1. Identify best practices to ensure that knowledge of operations is captured in the user interfaces (e.g., users are not able to enter parameters that are out of bounds).
2. Develop process for tool verification and validation, certification, re-verification, revalidation, and recertification based on operational data and trending.
3. Develop standard for documentation, configuration management, and quality assurance.
4. Identify any training or certification requirements to ensure proper operational capabilities.
5. Provide a plan for tool management, maintenance, and obsolescence consistent with modeling/simulation environments and the aging or changing of the modeled platform or system.
6. Develop a process for user feedback when results appear unrealistic or defy explanation.”

The implementation plan was developed by the Development Team at LaRC, led by Mr. Thomas Zang. The plan was approved by the OCE on April 19, 2005, and fiscal year (FY) 05 funding was provided by the OCE in May 2005.


In September 2005, the Program Analysis and Evaluation Office reviewed all the Diaz Actions and recommended that Diaz Action #4 be completed.

On October 23, 2005, Mr. Gregory Robinson requested the NASA Engineering and Safety Center (NESC) to assume sponsorship and oversight of this activity. This assessment officially began on January 13, 2006, under an informal plan from the Development Team. The formal plan for the FY 07–08 activities was developed by the Topic Working Group, which was formed in February–June 2006 under the auspices of the NASA Technical Standards Working Group. The NESC Review Board (NRB) approved this formal plan on March 29, 2007. The final report was presented to the NRB on November 20, 2008.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 2 of 112


2.0 Signature Page

_____ Dr. Thomas A. Zang Date	_____ Maria Babula Date
_____ Dr. Steve R. Blattnig Date	_____ Mr. William J. Bertch Date
_____ Mr. Richard E. Davis Date	_____ Mr. Lawrence L. Green Date
_____ Mr. Joseph P. Hale Date	_____ Dr. Michael J. Hensch Date
_____ Dr. James M. Luckring Date	_____ Dr. Unmeel B. Mehta Date
_____ Mr. Joseph H. Morrison Date	_____ Mr. Gary E. Mosier Date
_____ Dr. Ram K. Tripathi Date	_____ Dr. Martin J. Steele Date
_____ Mr. Andre J. Sylvester Date	_____ Mr. Jody Woods Date

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 3 of 112

3.0 Team List

Name	Discipline	Organization/Location
Core Team		
Thomas A. Zang	Development Team Lead & Topic Working Group Chair	LaRC
Steve R. Blattnig	Development Team member	LaRC
Lawrence L. Green	Development Team member	LaRC
Michael J. Hemsch	Development Team member	LaRC
James M. Luckring	Development Team member	LaRC
Joseph H. Morrison	Development Team member	LaRC
Ram K. Tripathi	Development Team member	LaRC
Maria Babula	Topic Working Group member	GRC
William J. Bertch	Topic Working Group member	JPL
Richard E. Davis	Topic Working Group member (prior to March 31, 2007)	LaRC
Lawrence L. Green	Topic Working Group member (after April 1, 2007)	LaRC
Joseph P. Hale	Topic Working Group member	MSFC
Unmeel B. Mehta	Topic Working Group member	ARC
Gary E. Mosier	Topic Working Group member	GSFC
Martin J. Steele	Topic Working Group member	KSC
Andre J. Sylvester	Topic Working Group member	JSC
Jody Woods	Topic Working Group member	SSC
Program/Project/Organization Liaisons		
Michael P. Blythe	OCE Liaison (prior to May 18, 2006)	HQ
Harold M. Bell	OCE Liaison (after May 19, 2006)	HQ
Dawn M. Schaible	NESC Liaison	NESC
Kenneth L. Johnson	NESC Topic Working Group Representative	NESC
Administrative Support Personnel		
Eric Pope	Technical Writer	LaRC/ATK
Charles V. Dunton	Workshop Facilitator	LaRC

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 4 of 112

3.1 Acknowledgements

The following NASA employees served as interim or alternate members of the Topic Working Group: Jeffrey Estefan (JPL), Robert Field (SSC), Richard Halverstad (JPL), Joseph Hubbard (JSC), Thomas McCarthy (GSFC), Galen Overstreet (JSC), Jeffrey Rusick (GRC), Harry M. Ryan (SSC), Nina Scheller (ARC), and Stephen Wall (JPL). Technical support for this activity was provided primarily by the National Institute for Aerospace (NIA), Aegis Technologies, and Sandia National Laboratories. The NIA facilitated consulting services from Osman Balci (Virginia Tech), Scott Harmon (Zetetix), Sankharan Mahadevan (Vanderbilt University), and David Nicol (University of Illinois). Aegis employees who consulted on this activity were Alleen Bray, Lisa Caine, Bobby Hartway, Danny Thomas, and William Waite. Audrey Milroy of QTEC also consulted. Sandia National Laboratories, Albuquerque consultants were William Oberkamp, David Percy, Martin Pilch, and Timothy Trucano. Simone Youngblood of the (then) Defense Modeling and Simulation Office hosted a workshop in May 2006 that was related to scales for assessing models and simulations (M&S).


The following participated in the First Scales Workshop, March 6-7, 2007: Scott Harmon, Hans Mair (Institute for Defense Analysis), William Oberkamp, Kevin Tucker (MSFC), William Waite, and Simone Youngblood.

The following Space Shuttle Orbiter Reinforced Carbon-Carbon (RCC) Return to Flight (RTF) Impact Damage Threshold Assessment Team provided important background and feedback information in the early stages of the M&S Standard development: LaRC (Karen Lyle, Edwin Fasanella) and GRC (Matthew Melis, Kelly Carney, Mike Pereira and Robert Goldberg) members of

In addition to many of the consultants named above, the following NASA personnel provided comments on the first two versions of the M&S Standard: Natalia Alexandrov, Karen Lyle, Peter Gnoffo, and Neal Zapp. Numerous NASA individuals participated in the pilot studies in Spring 2007, and dozens of NASA individuals provided formal comments on the Interim M&S Standard during Summer 2007.

Mr. James Kelly, pilot of STS-114, provided his perspective on the weaknesses of the reporting of the many results from M&S that he observed during the formal reviews leading up to that flight. Mr. Charles Daniel (Valador), a co-author of the Space Shuttle RTF Task Group Final Report [ref. 2], provided helpful background on the comments on M&S provided in the report's Annex A.2.

Initial OCE oversight of this activity was provided by Mr. Michael Blythe, and initial NESC oversight was provided by Mr. John Stadler.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 5 of 112

4.0 Executive Summary


After the Columbia Accident Investigation Board (CAIB) report [ref. 3], the NASA Administrator *at that time* chartered an executive team (known as the Diaz Team) to identify the CAIB report elements with Agency-wide applicability, and to develop corrective measures to address each element. The following report documents the chronological development and release of an Agency-wide Standard for Models and Simulations (M&S) (NASA Standard 7009) in response to Action #4 from the report, “*A Renewed Commitment to Excellence: An Assessment of the NASA Agency-wide Applicability of the Columbia Accident Investigation Board Report, January 30, 2004*” [ref. 1]. This action was to “Develop a standard for the development, documentation, and operation of models and simulations”. The NASA Chief Engineer *at that time* augmented the detailed description of this action in a memo to the NASA Engineering Management Board (EMB) dated September 1, 2006 (Appendix A). The major addition to the objectives of Diaz Action #4 was the inclusion of the standard method to assess the credibility of M&S. This is referred to as the “credibility assessment scale.”

The first part of this action—the development of a draft standard for M&S—was accomplished by the Development Team. After some modifications by the Topic Working Group, this draft was issued as the NASA Interim M&S Standard for M&S, NASA-STD-(I)-7009 [ref. 4] on December 1, 2006. Subsequently, the Topic Working Group made extensive revisions to the credibility assessment scale and then to the M&S Standard itself in response to the NASA-wide review. Their product was the Revised M&S Standard, which was delivered to the NASA Technical Standards Program Office on November 16, 2007. This underwent additional revisions as a result of the EMB review, and that final document became the Permanent NASA Standard for M&S, NASA-STD-7009 [ref. 5] in July 2008.

In the course of the Interim, Revised, and Permanent M&S Standards development, review, and release, a number of Development Team and Topic Working Group findings, observations, recommendations, alternate viewpoints, and lessons learned were identified. These are directed to the OCE and located in Sections 9.0, 10.0, and 12.0.

The following Development Team and Topic Working Group recommendations are made:

- R-1. NASA should integrate the M&S Standard into the NASA guidance hierarchy.**
- R-2. NASA should coordinate with other organizations and professional societies to further mature the M&S Standard.**
- R-3. NASA should sponsor development of Recommended Practices Guides.**
- R-4. NASA should re-assess the requirements on recommended practices that were removed from the Interim M&S Standard.**

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 6 of 112

- R-5. NASA should refine how submodels are treated in the credibility assessment scale.**
- R-6. Information regarding credibility assessment scale usage should be collected to determine effectiveness and provide data for further revision.**
- R-7. NASA should clarify the operational meaning of the term “consensus” in NASA Interim Directive on Interim NASA Technical Standards.**


Topic Working Group Recommendations:

- R-8. NASA should sponsor the development of Recommended Practices Guides along disciplinary lines. This responsibility might best be delegated to the NASA Technical Fellows.**
- R-9. NASA should collect data on the scope decisions, the cost impact, and the credibility assessment scale usage of the M&S Standard.**
- R-10. NASA should develop, by application domain, an M&S “validation lessons learned” database.**
- R-11. An NPD and/or NPR should call out the M&S Standard.**
- R-12. Centers should share with each other their plans and other guidance for implementation of the M&S Standard.**

5.0 Assessment Plan

This assessment consisted of two phases, with the major activities illustrated in [Figures 5.0-1 and 5.0-2](#). In Phase 1 (May 2005 - August 2006), the Development Team from LaRC performed background research, formulated the general approach, and developed the first three versions of the M&S Standard. An informal review of the second draft was conducted which solicited comments from the Centers. In Phase 2 (August 2006 - November 2007), the Topic Working Group, with membership from all Centers except for DFRC revised Version 3 into Version 4, which became the Interim NASA M&S Standard for M&S; oversaw the roll-out of the M&S Standard at their Centers; fostered the pilot studies and formal comments at their Centers; and revised the Interim M&S Standard into the version submitted for EMB approval as Version 5. A substantial part of this work involved consensus on a single credibility assessment scale.

There were three principal documents produced during this task. The term *Interim M&S Standard* is used herein to refer to the Interim NASA Standard for Models and Simulations [ref. 4], and discussed in Section 7.3. The term *Revised M&S Standard* refers to Version 5 approved by the Topic Working Group in response to the Agency-wide review discussed in Section 7.5. The term *Permanent M&S Standard* refers to the subsequent version [ref. 5] formally issued in July 2008 that included changes resulting from the EMB review discussed in Section 7.6.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 7 of 112

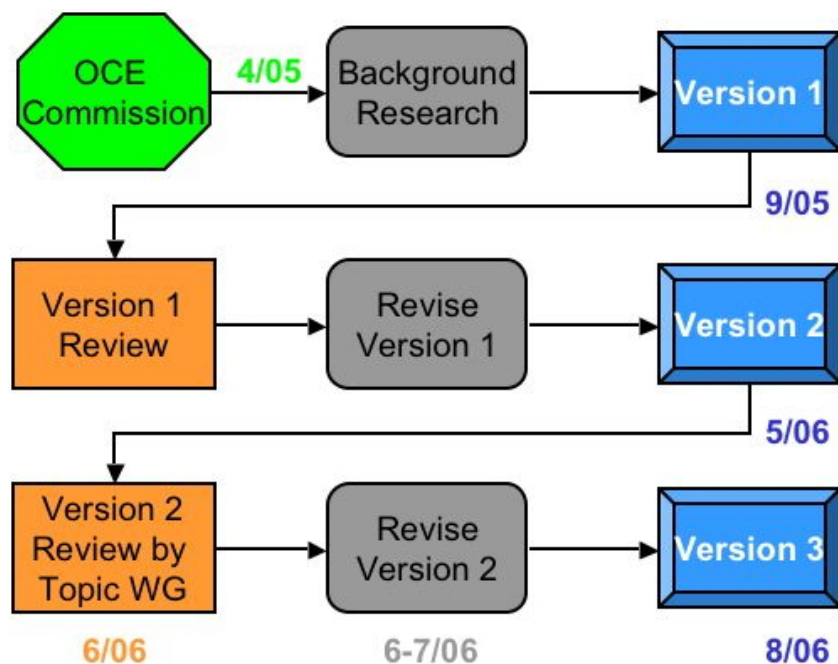


Figure 5.0-1. Phase 1 of the M&S Standard Development

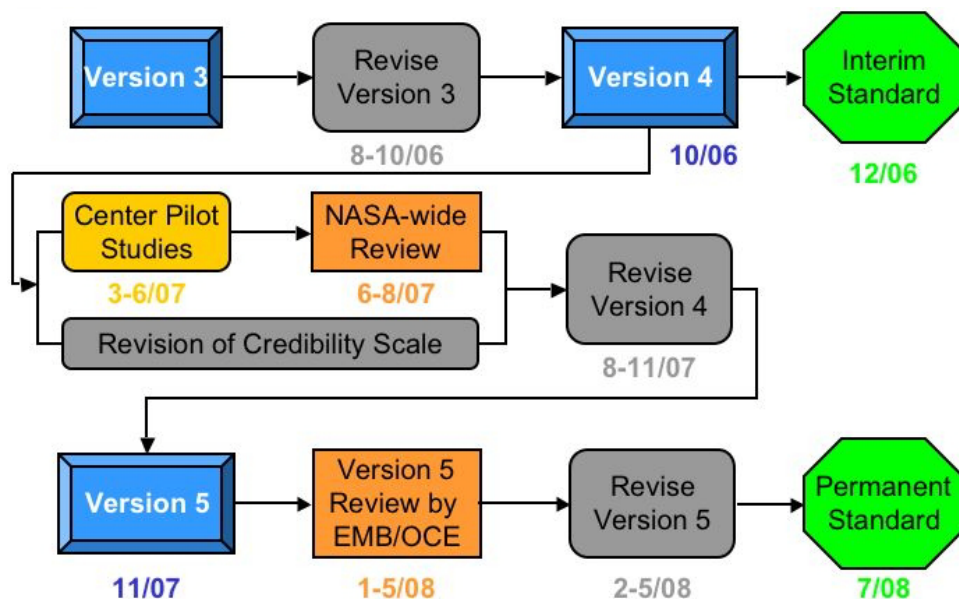



Figure 5.0-2. Phase 2 of the M&S Standard Development

Phase 1 made extensive use of external consultants from the Department of Energy (DoE), the Department of Defense (DoD), NASA contractors, and academia to determine the state-of-the-

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 8 of 112

practice in M&S guidance (including, but not limited to, formal standards), and to solicit comments on the first two versions of the M&S Standard. Phase 2 utilized weekly web-based meetings and five face-to-face workshops to resolve key issues on the scale and the disposition of formal comments. The first three of these workshops utilized the services of a trained facilitator, which proved useful in focusing and documenting the discussions and decisions.

Decision-making by the Topic Working Group in Phase 2 was more formalized than it was by the Development Team in Phase 1. The quorum of voting members required for a Topic Working Group decision was 6 (of 9) voting members. The Topic Working Group made most of its decisions by a supermajority rule (e.g. if 8 voting members were present, then a 6-2 vote was decisive, but a 5-3 (or 4-4) vote required additional discussion). Furthermore, the Topic Working Group eventually added rules to curb the temptation to revisit previous formal decisions: (1) a formal motion plus a second was required even to begin discussion of a previous decision, and (2) overturning a previous decision required a supermajority in favor of the overturn.

The original plan submitted by the Development Team called for submittal of a draft M&S Standard to OCE by April 2006 (later extended to July 2006). This draft would then undergo the normal review process for NASA Standards. In June 2006, the NASA Chief Engineer determined that the M&S Standard should be issued as a NASA interim standard in Fall 2006, and then undergo the normal review process. Issuance as an interim M&S Standard required consensus approval by a Topic Working Group under the auspices of the NASA Technical Standards Working Group. Phase 1 ended when the Development Team submitted Version 3 as their final deliverable in August 2006. Phase 2 commenced with the Topic Working Group review of Version 3 that same month.

(In this report, the term *NASA Chief Engineer* is only used to refer to the individual in that position, whereas the term *Office of the Chief Engineer* refers generically to those individuals in the OCE with direct oversight responsibility for this activity.)


6.0 Problem Description and Scope

This section summarizes the written guidance provided for the development of the M&S Standard.

6.1 Columbia Accident Investigation Board

The CAIB report [ref. 3] contained four findings that were relevant to the development of the M&S Standard:

“F.6.3-10: The Team’s assessment of possible tile damage was performed using an impact simulation that was well outside Crater’s test database. The Boeing analyst was inexperienced in the use of Crater and the interpretation of its results. Engineers with ex-

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 9 of 112

tensive Thermal Protection System expertise at Huntington Beach were not actively involved in determining if the Crater results were properly interpreted.”

“F.6.3-11: Crater initially predicted tile damage deeper than the actual tile depth, but engineers used their judgment to conclude that damage would not penetrate the densified layer of tile. Similarly, RCC damage conclusions were based primarily on judgment and experience rather than analysis.”

“F6-3-13: The assumptions (and their uncertainties) used in the analysis were never presented or discussed in full to either the Mission Evaluation Room or the Mission Management Team.”

“F.10.1-4: The FAA and U.S. space launch ranges have safety standards designed to ensure that the general public is exposed to less than a one-in-a-million chance of serious injury from the operation of space launch vehicles and unmanned aircraft.”

These findings led to the Space Shuttle Program-specific recommendation:


“R3.8-2: Develop, validate, and maintain physics-based computer models to evaluate Thermal Protection System damage from debris impacts. These tools should provide realistic and timely estimates of any impact damage from possible debris from any source that may ultimately impact the Orbiter. Establish impact damage thresholds that trigger responsive corrective action, such as on-orbit inspection and repair, when indicated.”

6.2 A Renewed Commitment to Excellence: An Assessment of the NASA Agency-wide Applicability of the Columbia Accident Investigation Board Report

In its general discussion of these issues, the *A Renewed Commitment to Excellence: An Assessment of the NASA Agency-wide Applicability of the Columbia Accident Investigation Board Report*, hereafter referred to as the “Diaz Team report” [ref. 1] suggested:

“All programs should produce, maintain, and validate models to assess the state of their systems and components. These models should be continually updated and validated against experimental and operational data to determine appropriate courses of action and repair. The value of the models should be assessed with respect to their ability to support decision making in a timely way so as not to lead the decision maker to a conflict between costly action versus effective action in the interest of safety or mission success.”

“Personnel need to be adequately trained in model use, limitations, and escalation procedures when issues arise. Engineers, when faced with results that defy “reality checks,” should double check the model then raise their concerns.”

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 10 of 112

“NASA policies recognize requirements for public safety. Those policies should be reviewed and the models used should be continually updated and assessed with respect to value in supporting timely decision making.”

The detailed statement of Diaz Action #4 [ref. 1] was


“Develop a standard for the development, documentation, and operation of models and simulations.

1. Identify best practices to ensure that knowledge of operations is captured in the user interfaces (e.g., users are not able to enter parameters that are out of bounds).
2. Develop process for tool verification and validation, certification, re-verification, revalidation, and recertification based on operational data and trending.
3. Develop standard for documentation, configuration management, and quality assurance.
4. Identify any training or certification requirements to ensure proper operational capabilities.
5. Provide a plan for tool management, maintenance, and obsolescence consistent with modeling/simulation environments and the aging or changing of the modeled platform or system.
6. Develop a process for user feedback when results appear unrealistic or defy explanation.”

6.3 NASA Chief Engineer Guidance

The specific goals stated for the M&S Standard in a memo by the NASA Chief Engineer dated September 1, 2006 (see Appendix A), were that the M&S Standard will

- Ensure that the credibility of M&S results is properly conveyed to those making critical decisions.
- Assure that the credibility of M&S meets the project requirements.
- Establish M&S requirements and recommendations that will form a strong foundation for disciplined (structure, management, control) development, validation and use of M&S within NASA and its contractor community.
- Include a standard method to assess the credibility of the M&S presented to the decision-maker when making critical decisions (i.e., decisions that effect human safety or mission success) using results from M&S.
- Establish a common set of terms and a uniform way for M&S practitioners to communicate the credibility of M&S.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 11 of 112

- Be responsive to Diaz Action #4.”

The scope of the M&S Standard was based on guidance from the OCE. The general statement of the scope provided in June 2005 was that the M&S Standard should apply to those M&S whose results are used for decisions that may impact human safety or mission success. A companion decision made by the OCE *at that time* was that the M&S Standard should not apply to software used for control systems and displays. The scope was refined over the course of the development, culminating in the specific language in the Permanent M&S Standard: “This standard applies to M&S used by NASA and its contractors for critical decisions in design, development, manufacturing, ground operations, and flight operations. This standard also applies to use of legacy as well as commercial-off-the-shelf (COTS), government-off-the-shelf (GOTS) and modified-off-the-shelf (MOTS) M&S to support critical decisions. ... This standard does not apply to M&S that are embedded in control software, emulation software, and stimulation environments.” The key phrase “critical decisions” is explained thusly in the Permanent M&S Standard: “Critical decisions based on M&S results, as defined by this standard, are those technical decisions related to design, development, manufacturing, ground, or flight operations that may impact human safety or program/project-defined mission success criteria.” Furthermore, the Permanent M&S Standard includes a risk assessment process for use by the Program/Project and the technical authority in their determination of which M&S are in scope.

Requirements on the *presentation* is a rather unique approach for a standard, but gets to the heart of the issues raised in the CAIB report [ref. 3], which emphasized that key information was not properly conveyed to the decision-makers.

7.0 Major Activities


This section discusses the major activities by the Development Team, the Topic Working Group, the Technical Standards Working Group, the OCE, and the EMB during the 3-year process of the development, reviews, and revisions of the M&S Standard. The discussion is grouped by topic rather than by the timeline.

The M&S Standard evolved through six distinct versions. The Development Team produced Versions 1- 3 and the Topic Working Group was responsible for Versions 4 and 5. Version 6 consisted of the modifications to Version 5 that resulted from the EMB review (Section 7.6).

7.1 Timeline of M&S Standard Development

[Table 7.1-1](#) summarizes the noteworthy activities during Phase 1 (see also [Figure 5.0-1](#)), and [Table 7.1-2](#) summarizes the noteworthy activities during Phase 2 (see also [Figure 5.0-2](#)). The third column indicates the section(s) in which the activity is discussed.


Table 7.1-1. M&S Standard Phase 1 Timeline

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 12 of 112

Date	Activity	Section
April 2005	OCE charts Development Team to respond to Diaz Action #4	1.0
April–June 2005	Development Team reviews existing NASA and non-NASA guidance documents	7.2.1
June–July 2005	Space Shuttle Orbiter RCC RTF Impact Damage Threshold Assessment Team pilot study	7.2.2
July 2005	Development Team completes uncertainty structure table	7.2.3
Aug.–Sept. 2005	Development Team writes Version 1	7.2.4
September 2005	Program Analysis & Evaluation Office recommends completion of Diaz Action #4	7.2.1
September 2005	Development Team releases Version 1 to OCE and consultants	7.2.4
October 2005	Status review by OCE	7.2.1
Nov. 2005–April 2006	Development Team revises Version 1	7.2.4
Feb.–June 2006	M&S Standard Topic Working Group formed	7.2.4
March 2006	NASA Chief Engineer suggests inclusion of a Scale	7.2.3
April 2006	Development Team releases preliminary draft of Version 2 to consultants	7.2.4
May 2006	Development Team releases Version 2 to Topic Working Group	7.2.4
May 2006	NASA / DoE /DoD meeting at DMSO (scales / maturity matrices)	7.2.3
May–July 2006	Briefings at ARC, GSFC, GRC, JPL, JSC, KSC, and MSFC	7.2.4
June 2006	Topic Working Group reviews Version 2	7.2.4
June 2006	NASA Chief Engineer targets the M&S Standard for release as a NASA interim standard	7.2.3
July 2006	NASA Chief Engineer directs inclusion of a credibility assessment scale in the M&S Standard	7.2.3
June–July 2006	Development Team revises Version 2 and adds first credibility assessment scale	7.2.4
August 2006	Development Team releases Version 3 to Topic Working Group and OCE as final team deliverable	7.2.4

Table 7.1-2. M&S Standard Phase 2 Timeline

Date	Activity	Section
------	----------	---------


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 13 of 112

Date	Activity	Section
August 2006	Topic Working Group meeting at HQ	7.3.2
Aug.–Sept. 2006	Topic Working Group revises Version 3	7.3.2
September 2006	Topic Working Group meeting at JSC	7.3.2
October 2006	Topic Working Group approves Version 4 and releases to TSPO & OCE	7.3.2
December 2006	OCE issues Interim M&S Standard	7.3.2
Feb.–March 2007	Topic Working Group rolls out Interim M&S Standard at Centers	7.3.3
March 2007	Scale Workshop 1 at LaRC	7.4.1
Mar.–June 2007	Centers conduct pilot studies	7.3.4
April 2007	Scale Workshop 2 at JSC	7.4.2
May 2007	Scale Workshop 3 at KSC	7.4.3
May–Aug. 2007	Agency-wide review of the Interim M&S Standard	7.5
July 2007	NASA / DoE workshop at Sandia National Laboratories	7.4.5
August 2007	Scale Workshop 4 at GSFC	7.4.4
Aug.–Nov. 2007	Topic Working Group disposition of formal comments	7.5.1
September 2007	Comment disposition meeting at JPL	7.5.1
November 2007	Topic Working Group approves Version 5 (Revised M&S Standard) and releases to TSPO and OCE as final Topic Working Group deliverable	7.5.1
Jan.–May 2008	Version 5 review by EMB/OCE	7.6
May 2008	Center objections discussed at EMB Meeting at MSFC	7.6
March–May 2008	Revision of Version 5 per EMB/OCE review	7.6
May 2008	Version 6 (Permanent M&S Standard) released to OCE	7.6
July 2008	OCE issues Permanent M&S Standard	7.6

7.2 Phase 1 Activities

7.2.1 Review of Existing M&S Guidance and Standards

During the May–July 2005 timeframe, the Development Team conducted an extensive search for relevant M&S guidance and standards in NASA and other Federal Agencies. The initial search identified by title approximately 300 documents that appeared relevant. These were downselected to 100 based on a review of the document abstracts. These final 100 documents

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 14 of 112

were each reviewed in depth by two members of the Development Team. The countless publications by individuals or small groups on M&S guidance were not considered germane for assessing the state of consensus standards or guidance for M&S.

The following findings resulted from this review:

- F-1. Current NASA guidance is oriented towards control systems and displays. Quality assurance and configuration management are well covered, but the unique, critical aspects of M&S are not addressed (i.e., validation against experimental or flight data, and uncertainty quantification).**
- F-2. No Federal Agency has an M&S standard, although the DoD has extensive M&S guidance, and the Nuclear Regulatory Commission has standards for control systems and displays.**
- F-3. Relevant M&S guidance is strongly focused on the development phase of the M&S life-cycle, and especially upon verification and validation. There is little guidance on the operations of M&S and virtually no guidance on the maintenance of M&S.**


The most relevant existing guidance on M&S includes the DoD's VV&A [Verification, Validation and Accreditation] Recommended Practices Guide [ref. 6], the American Institute of Aeronautics and Astronautics' (AIAA) Guide for Verification and Validation of Computational Fluid Dynamics Simulation [ref. 7], the American Society of Mechanical Engineers' (ASME) Guide for Verification and Validation in Computational Solid Mechanics [ref. 8], and Sandia National Laboratories' Concepts for Stockpile Computing [ref. 9].

- F-4. NASA has no policy or procedural requirements for M&S, except for the software engineering aspects covered by NPD 2820.1B, NASA Software Policies, and NPR 7150.2, NASA Software Engineering Requirements.**

In July 2005, the report of the Space Shuttle RTF Group [ref. 2] was issued. Annex A.2 contained many remarks that reinforced the need for an M&S standard. Relevant excerpts are provided in Appendix B of this report.

The Program Analysis and Evaluation Office reviewed all open Diaz Actions in September 2005. The following information on Diaz Action #4 was provided from that review:

- “Applicability
 - A general M&S standard would simplify the development of consistent, discipline-specific standards by the Technical Warrant Holders
- Continuing Value
 - Annex A.2 of the Stafford-Covey Return to Flight Task Group report [ref. 2] emphasizes the need for a NASA M&S Standard

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 15 of 112

- Related or Overlapping Activities
 - NASA has existing or imminent NPDs, NPRs and standards that cover many of the generic software engineering aspects of the Diaz #4 requirements (e.g., configuration management and quality assurance)
 - However, the unique, critical aspects of Models and Simulations (M&S) are not addressed by existing NASA or NASA-preferred documents, especially validation against experimental or flight data & uncertainty quantification”

As a result of its review, the Program Analysis and Evaluation Office recommended that Diaz Action #4 be completed.

In discussions with the OCE sponsor in August 2005, the Development Team recommended, and the OCE concurred, that Diaz Action #4 should focus on the development of a Standard (as opposed to an NPR or Guidebook), and that this standard should consist of explicit requirements and recommendations.


In the October 24, 2005 meeting in which the FY 05 results were reported, the Development Team recommended that the M&S Standard be anchored to a NPD or NPR. OCE indicated that they favored covering the M&S Standard in a revision to NPR 7123.1, Systems Engineering Processes and Requirements.

At the meeting the OCE concurred with the Development Team recommendation that the M&S Standard address “what” shall (requirement) or should (recommendation) be done. It would address neither “who” is responsible (an NPR-level issue), nor “how” it should be done (a Guideline-level issue). The domain of M&S is so broad that how requirements are implemented will be application-specific.

7.2.2 Space Shuttle Orbiter RCC Impact Damage Threshold Assessment Team Pilot Study

The initial plan submitted in April 2005 called for several pilot studies to be conducted. Due to the time constraints only one pilot study was performed—with the LaRC and GRC members of the Space Shuttle Orbiter RTF RCC Impact Damage Threshold Assessment Team. This team was responsible for the NASA response to the RCC leading edge portion of the CAIB Recommendation R3.8.2 (Section 6.1). An overview of this M&S project is given in ref. 10. This was the single M&S activity most relevant to the CAIB recommendation that led to Diaz Action #4.

This pilot was conducted in June–July 2005. The Development Team met face-to-face with the team lead, Ed Fasanella of LaRC, and with other civil servant team members: Karen Lyle of LaRC, and Matt Melis (GRC Team Lead), Kelly Carney, Robert Goldberg, Brad Lurch, Mike

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 16 of 112

Pereira, and Duane Revilock of GRC. During these meetings, which took approximately 3 days, the pilot team thoroughly explained the processes followed in conducting experiments, developing material models, developing the full computational model (using LS-DYNA™), and undergoing technical reviews. These discussions influenced some of the choices in Version 1. Members of this team provided detailed comments on Versions 1 and 2. The LaRC team members also provided feedback on the two scales in the Interim M&S Standard based on application of both to their M&S project.

7.2.3 M&S Scales/Maturity Matrices


The first indication that OCE was interested in having what has come to be called the “credibility assessment scale” came during a meeting between the team lead and the NASA Chief Engineer on March 1, 2006. The Chief Engineer stated that he wanted a level of rigor scale in the M&S Standard and not in any supporting document such as an M&S Guidebook. At this meeting the only scale that the team lead had at hand was the Uncertainty Structure [ref. 11]. The Chief Engineer indicated that something along the lines of this matrix was what he had envisioned, albeit one covering all aspects of M&S and not just uncertainty.

The Defense Modeling and Simulation Office (DMSO) hosted a small workshop on M&S maturity matrices on May 9-10, 2006, at which the Validation Process Maturity Model (VPMM) [ref. 12], the Predictive Capability Maturity Model (PCMM) [ref. 13], the Uncertainty Structure scale [ref. 11], and the Simulation Readiness Level (SRL) scale [ref. 14] were presented and critiqued. Each of these scales had been developed by a small, homogeneous group that consisted of individuals from at most 2 branches at a single laboratory (in NASA terms). Each scale had evolved over a period of 2-4 years and each was still evolving. All agreed that there had been arguments even among their small group on the factors and level definitions for their scale. All agreed that application of the scales to a broad spectrum of actual M&S activities was essential to development of a useful scale.

The Development Team reached the following conclusions:

- F-5. There does not presently exist an M&S scale with the specific objectives desired for the M&S Standard.**
- O-1. Development of a rigor scale is extremely difficult, even for a small, homogeneous group, and even restricted to M&S using mathematical models based on partial differential equations (PDEs).**

A second meeting was held with the NASA Chief Engineer on July 19, 2006. The Development Team proposed the use of an “adherence scale” (i.e., how well were the M&S Standard requirements satisfied in lieu of a “rigor scale”) because of the difficulty in devising an acceptable rigor scale that is applicable to all types of M&S. The NASA Chief Engineer directed the Development Team to produce a rigor scale (called the credibility assessment scale in the

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 17 of 112

Permanent M&S Standard), even if is very high level and even if it had to be restricted to just PDE-based M&S. At this same meeting the NASA Chief Engineer also indicated that since he had determined that the M&S Standard should be issued as a NASA interim standard (see Sections 5.0 and 7.3.1), he desired that Version 3 be completed by mid-August 2006.

As a result of this guidance, over the following three weeks the Development Team devised the rigor scale that appeared in Appendix A2 of Version 3 (and ultimately in the Interim M&S Standard). The detailed rationale behind this scale is available in Reference 15. The context of this scale was that the credibility of the results is determined by the decision-maker based upon two pieces of information:

1. The estimate of the uncertainty in the results (Req. 4.8.3 of the Interim M&S Standard)
2. The objective assessment of the rigor of the processes used to generate the results (including the uncertainty estimate) (Req. 4.8.5 of the Interim M&S Standard)

The Development Team's perspective was that

O-2. The credibility scale is not a stand-alone assessment of factors influencing credibility, but rather the credibility assessment scale plus the uncertainty statement combine to influence the credibility assessment by the decision-maker.

Subsequently, the MSFC member of the Topic Working Group proposed an alternative credibility assessment scale that appeared in Appendix A3 of the Interim M&S Standard. See reference 16 for details on this scale.


7.2.4 First Three Versions of the Standard

The Development Team produced three versions of the M&S Standard. This section discusses the major characteristics of those versions and the results of the reviews that were performed. The Topic Working Group produced an additional two versions in Phase 2. Those are covered in Sections 7.3.2 and 7.5. The final version, which contains changes made as a result of the EMB review, is discussed in Section 7.6.

Version 1

Version 1 was the straw man intended to: (1) satisfy the OCE need for a specific deliverable at the end of FY 05, and (2) provoke comments from reviewers to assist the Development Team in sharpening their thinking on the goal, objectives, and structure of the M&S Standard.

At this point (September 2005) in the development process the vision of the Development Team was that

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 18 of 112

1. The M&S Standard should require the NASA M&S development and operations communities to report their processes in such a manner that the decision-maker can **quantitatively** assess the associated risk for safety and mission assurance.
2. This standard should be supplemented by a recommended practices guide that enables the above.
3. The M&S Standard should be such that working troops and project managers can understand and accept its processes.

The objectives for this standard were “to make available standard practices for assuring and quantitatively assessing the quality M&S results throughout their development and use for specific applications, together with timely and complete reporting of the quality assurance processes and assessments.”

The key aspects were:

- Documentation and reporting, etc.
- Defensible confidence building
- Defensible uncertainty quantification


Version 1 of the M&S Standard, entitled “Quality Assurance for Models and Simulations”, was circulated on September 21, 2005. Comments were submitted by Osman Balci, Robert Gravitz, Sankharan Mahadevan, Audrey Milroy, David Percy, William Oberkampf, David O’Neil, and David Schuster. The comments led to a complete re-thinking of the M&S Standard during Fall 2005.

This is the version that was discussed at the status update to the OCE on October 24, 2005 of the FY 05 accomplishments.

Version 2

Two of the major objections to Version 1 were its emphasis on quality, and the large number of requirements. The Development Team concluded that couching the M&S Standard in terms of quality would impede the acceptance of the M&S Standard by practitioners. In addition to this philosophical change, the principles articulated in November 2005 that guided the development of Version 2 were

1. Keep it simple and concise.
2. Stick to the essence.
3. The prize goes to the shortest, clearest document.
4. Don’t have a requirement that will frequently be waived.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 19 of 112

5. Ideally, future debates about revisions to the M&S Standard should be about putting additional requirements or recommendations into the M&S Standard, and not about taking existing requirements or recommendations out of the M&S Standard.

The ground rule that the Development Team adopted for the inclusion of a requirement in the M&S Standard was that the team had to agree unanimously on its inclusion. This rule helped to enforce the principles listed above.

As Version 2 matured, the specific objectives of that standard were articulated as “The requirements and recommendations of the M&S Standard are intended to assure that

- Decision-makers can assess the credibility of results from M&S.
- Violations of the limitations of the M&S are apparent to decision-makers, and a summary of the limitations are easily accessible to the decision-maker.
- Processes for modeling and simulation are transparent to decision-makers.
- Rigor of the M&S can be evaluated against the program or project requirements.
- Results from the M&S are reproducible by M&S domain experts”.

The Development Team’s experiences in devising the first two versions of the M&S Standard are summarized as


O-3. Frequent face-to-face meetings were essential for the initial formulation of the M&S Standard.

A preliminary draft of the second version of the M&S Standard was circulated on March 31, 2006. Their comments led to a significant refinement of the final Version 2 and, in particular, to the articulation of two high-level goals:

1. “The primary goal of this standard is to ensure that the credibility of the results from models and simulations is properly conveyed to those making critical decisions based in full or in part upon the results of models and simulations.
2. The secondary goal is to increase that credibility.”

This version did not include a credibility assessment scale.

The Topic Working Group for the M&S Standard was constituted during February–June 2006. The Kick-Off Meeting for the Topic Working Group review of Version 3 of the M&S Standard occurred on May 23, 2006 at GRC. At this point, 8 of the 10 Centers were involved. JPL joined the Topic Working Group in June 2006, and DFRC chose not to participate. Centers present in person at the Kick-Off Meeting were ARC, GRC, and MSFC. JSC, GSFC, KSC, LaRC, and SSC participated via teleconference. The Topic Working Group was briefed on the background on the M&S Standard, on its overall philosophy, the top-level decisions, a survey of the

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 20 of 112

requirements, and the process for providing feedback. Subsequently, visits were made to KSC (May 31), JPL (June 13), ARC (June 14), JSC (June 15), GSFC (June 26), and MSFC (July 14) to brief the Topic Working Group representative and/or others at the Center, and to solicit feedback.

During the discussions with the Orbiter RCC Impact Damage Threshold Assessment Team and these meetings at the various Centers, the Development Team observed that

O-4. Many engineers and program managers at NASA are unaware of the intended hierarchy of the agency guidance documents.

O-5. Many engineers at NASA are unaware of standards that are relevant to their work.

Approximately 306 comments were submitted on Version 2 as a result of this first Topic Working Group review. Of these, approximately two-thirds came from Topic Working Group members.

Version 3

Version 3 had numerous detailed changes made as a result of these comments. Approximately half of the comments were accepted. Many of the rejected comments were objections to having a standard at all or to specific choices, such as scope, for which guidance had been provided by OCE. In addition, Version 3 included a credibility assessment scale in response to the NASA Chief Engineer's July 19, 2006 direction (Section 7.2.3).

Version 3 was submitted to the OCE and the Topic Working Group on August 15, 2006. This marked the final deliverable of the Development Team and the conclusion of Phase 1.


At this point, the Topic Working Group became the responsible body. They revised Version 3, formally approved the resulting Version 4 on September 28, and submitted it to OCE (Section 7.3.2).

7.3 The Interim M&S Standard

7.3.1 Background on Interim M&S Standards

In mid-2006 the OCE determined that there were some standards needed by Programs and Projects that could not wait for the lengthy review process required of NASA permanent standards. A NASA Interim Directive on Interim NASA Technical Standards was formally issued on August 24, 2006. The role of the Topic Working Group was stated as follows:

“The Chair of the Topic Working Group shall prepare

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 21 of 112

- a. A consensus version of the candidate document which bears a NASA-STD-(I)-xxxx designation and the following footnote on the cover and each page: *“This document represents the technical consensus of the developing group but does not yet have final NASA approval.”*
- b. Certification that Center representatives participating in the Topic Working Group have reached consensus on all substantive technical issues.”


The Technical Standards Executive (Richard Weinstein) explained that the term “consensus” meant “absence of a sustained, substantive technical objection.”

7.3.2 Version 3 Review

The review and revisions to Version 3 were performed by the Topic Working Group in August and September 2006. The process for interim standards required that all 300 of the comments on version 2 be formally dispositioned. The Topic Working Group reviewed the disposition by the Development Team of the comments on Version 2, and reviewed the new material in Section 4.7 and Appendix A of the standard on the credibility assessment scale. (Note: Although Version 3, Version 4 and the Interim M&S Standard used the term “credibility scale” and not the term “credibility assessment scale” that is used in Version 5 and the Permanent M&S Standard, the latter term is used in this document for consistency.)

The rules for disposition of individual comments were that in order to disposition a comment (a) the Topic Working Group representative from the commenter’s Center needed to be present, and (b) a supermajority of these present was needed to a decision on the comment. The Topic Working Group held two meetings (with some members connected via teleconference) as part of their review of Version 3. The first was held August 30–31 at NASA Headquarters (HQ). The second was held September 5-6 at JSC. As a result of this process, a number of the proposed dispositions were changed by the Topic Working Group.

At the end of the meeting at NASA HQ, the NASA Chief Engineer met with the Topic Working Group to discuss his goals for the credibility assessment scale. Two detailed alternatives to the credibility assessment scale in Version 3 were offered by Topic Working Group members, one by Joe Hale (MSFC) and one by Unmeel Mehta (ARC). A major outcome of the JSC meeting was the decision (by a 6-3 vote) to use a multi-dimensional rather than a one-dimensional scale. This left the credibility assessment scale in Version 3 and the Hale alternative as options. The Topic Working Group was unable, in the time available to produce the Interim M&S Standard, to come to supermajority agreement on a single credibility assessment scale. Therefore, the Interim M&S Standard contained two alternatives located in Appendices A2 and A3, referred to hereafter as the *A2 scale* and the *A3 scale*, respectively.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 22 of 112

The Topic Working Group wrestled considerably with the definition of consensus given in the NASA Interim Directive on Interim NASA Technical Standards, and concluded that

F-6. The definition of a Topic Working Group consensus in the NASA Interim Directive on Interim NASA Technical Standards is not sufficiently precise.

At the end of this process, Version 4 was voted upon by the Topic Working Group. Three options were available: (a) concur; (b) can live with it (but have some reservations); or (c) non-concur. The results of the vote are recorded in [Table 7.3-1](#). Version 4 was submitted to the NASA Technical Standards Program Office on October 10, 2006.


Table 7.3-3. Topic Working Group Vote on the Interim M&S Standard

Center	Representative	Vote
ARC	Unmeel Mehta	Non-concur
GRC	Jeffrey Rusick	Can live with it
GSFC	Thomas McCarthy	Concur
JPL	Jeffrey Estefan	Can live with it
JSC	Galen Overstreet	Can live with it
KSC	Martin Steele	Can live with it
LaRC	Richard Davis	Concur
MSFC	Joe Hale	Can live with it
SSC	Jody Woods	Concur

The objections of the ARC representative are provided in Appendix C. As a result of discussions with OCE, the ARC EMB member did not sustain this objection on the assurance that the ARC concerns would be addressed in the path forward towards a permanent standard. Version 4 was adopted as the Interim M&S Standard after OCE changed the specific term “certification” to the generic term “endorsement” in one requirement and several level definitions in the two scales.

The NASA Chief Engineer issued the Interim NASA Standard for M&S, NASA-STD-(I)-7009 [ref. 4], on December 1, 2006.

The issuance of the Interim M&S Standard completed the work of the initial Topic Working Group. Afterwards the Topic Working Group members from GRC, GSFC, JPL, and JSC were replaced due to other commitments. The LaRC Topic Working Group member was replaced upon his retirement in March 2007.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 23 of 112

7.3.3 Rollout of the Interim M&S Standard

The OCE gave the Topic Working Group members the responsibility for rolling out the Interim M&S Standard at their Centers. The Topic Working Group developed several briefings for this purpose.

M&S Standard Executive Briefing: This briefing was given to the Center EMB members during February–March 2007. The purpose was to acquaint them with the motivation, scope, development process, key requirements and the credibility assessment scale. Furthermore, their support was solicited in the following areas:

- Ensure that your Center is represented on the Topic Working Group by a knowledgeable individual who can devote 10-15 percent of his/her time to the various Topic Working Group tasks
- Support the roll-out of Interim M&S Standard at your Center
- Sponsor pilot studies of Interim M&S Standard
- Encourage broad input into the Topic Working Group process from your Center on the Interim M&S Standard


M&S Standard Practitioner Briefing: This briefing was typically given at branch meetings and M&S team meetings during February–April 2007. It contained more details on the motivation, scope, development process, key requirements, and the credibility assessment scale than the Executive Briefing. It also gave an overview of the objectives and format of the pilot studies (Section 7.3.4).

7.3.4 Pilot Studies on the Interim M&S Standard

During the first half of 2007, pilot studies using the Interim M&S Standard were conducted by M&S teams at most Centers. These are listed in [Table 7.3-2](#). Their purpose was to provide feedback to the Topic Working Group on practical experience with the Interim M&S Standard and to ensure that the Center comments during the subsequent formal review were informed by this practical experience. Collectively, these pilot studies included a reasonably broad spectrum of M&S types (mostly using mathematical models based on differential equations, but also including a discrete model and a geometry model) and applications.

Table 7.3-4. Pilot Studies Using the Interim M&S Standard

Center	Application	Discipline(s)
ARC	Aerodynamic database supporting a Crew Exploration Vehicle-like atmospheric re-entry capsule	Aerodynamics


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 24 of 112

GSFC	Model of a Fine Guidance Sensor for the James Webb Space Telescope	Guidance, Navigation, and Control (GN&C)
JPL	Thermal model of Mars Exploration Rover Cruise Stage	Thermal
JPL	Mars Exploration Rover Entry, Descent, and Landing Simulation	Many
JPL	Simulation of the Kepler telescope, emphasizing detection of planet transients around the host star	Science
JPL	Model of an oceanographic sensor	Science
KSC	Extend @TM -based discrete event simulation to assess readiness and launch availability for the Crew Launch Vehicle	Logistics
KSC	Matlab@TM -based discrete event simulation for interplanetary logistics in building up and sustaining a lunar outpost	Logistics
LaRC	Uncertainty analysis of historical hurricane data, in support of hurricane predictions	Uncertainty
MSFC	Crew Launch Vehicle system readiness and launch availability	Processing, maintenance
SSC	Computational Fluid Dynamics Methane Technology Testbed model of a rocket thruster	Aerodynamics, Combustion

The Topic Working Group had developed a briefing for the pilot teams to supplement the M&S Standard Practitioners Briefing. The guidance provided to the pilot teams in this briefing consisted of:

“Pilot Study (PS) Purpose

- Determine whether the implementation of NASA-STD-(I)-7009 can fulfill the stated goals (see slide #9)
- Collect practical M&S experience with the Credibility Scales from developers, analysts & decision-makers
- Determine if the resulting M&S product would be better from using the Standard
- Assess the cost-benefit tradeoff

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 25 of 112

- the cost of implementing the M&S Standard
- any benefit from the M&S Standard versus its cost

Pilot Studies: Ideal Characteristics


- The ensemble of Pilot Studies from all Centers should cover the full spectrum of types of M&S (e.g., PDE-based, discrete-event-simulation based, cost models, operations models, design models, fault-tree models)
- Pilot Studies should include and focus on M&S activities that were sufficiently mature that results were presented to decision makers
- Each Pilot Study should be manageable (addressable within 2 months)

Pilot Study Team Responsibilities

- Apply the M&S Standard to the M&S in your Pilot Study
- Assess how well the M&S Standard fulfills stated goals of M&S Standard
- Evaluate the credibility scale, using both methods in Appendix A
- Early input on scale needed by 4/12
- Assess benefit from applying the M&S Standard to your M&S
- Assess additional cost imposed by applying the M&S Standard to your M&S
- Submit a report of the pilot results to the Topic Working Group
- Work with your TSWG representative to incorporate your comments into the NASA-wide formal review of the M&S Standard

Pilot Study Report Outline

- Pilot Study Background
- Briefly describe the M&S Project
- Identify any Requirements not covered and why
- Response to Credibility Scale Questionnaire
- Overall Comments/Concerns
- Ease of Use (Std. & Scale)
- Achievement of Goals of Standard
- Cost-Benefit
- Mock Briefing to Decision-maker (desirable)

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 26 of 112

- Optional (these comments can be made during Formal Review)
 - Clarity/Understandability of the Document
 - Response to Individual Requirement Questionnaire”

Two types of questionnaires were submitted to the pilot teams. One focused on the credibility assessment scales that were in the Interim M&S Standard, and the other covered the M&S Standard as a whole. A second similar questionnaire inquired how well the goals of the Interim M&S Standard are satisfied and what is the cost impact of compliance with the M&S Standard.

The pilot studies began on February 1. The scale questionnaire was submitted by April 20 and reviewed during the Second Scale Workshop (Section 7.6.2), and the final reports, whose major component was the responses to the second questionnaire, were submitted by the end of June. The final reports were due in the middle of the formal NASA-wide review. Hence, the pilot teams had a strong basis of experience with the Interim M&S Standard on which to base their formal comments.

Both questionnaires are discussed in more detail in Section 11, along with a summary of the responses to the multiple-choice questions. Section 11 also contains a data summary collected on the estimate of the cost impact of the M&S Standard.

These pilot studies were based on the Interim M&S Standard and their relevance to the Revised M&S Standard is limited.

7.4 Phase 2 Scale Workshops


The Topic Working Group noted that

O-6. The credibility assessment scale is outside the formal Diaz Action #4.

Developing an appropriate single scale for the M&S Standard was a major issue that the Topic Working Group worked to resolve during Phase 2. Indeed, this activity absorbed most of the Topic Working Group effort from March through July 2007. Much of the work on the scale was done at four face-to-face workshops, each lasting a day-and-a-half. A trained facilitator (Charles Dunton from LaRC) supported the first three workshops.

Detailed notes from these workshops are available on the team website. The most important considerations are summarized in the following sections.

A high-level summary of the process used in developing the final Scale is provided in [Figure 7.4-3](#). It started from a search of the literature, was influenced by Topic Working Group interviews of decision-makers, continued with the distillation of (orthogonal) factors, moved to selecting the final subset of key factors, and closed with determining the mechanism for

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 27 of 112

reporting the results of the assessment. More details of these steps are provided in the remainder of this subsection.

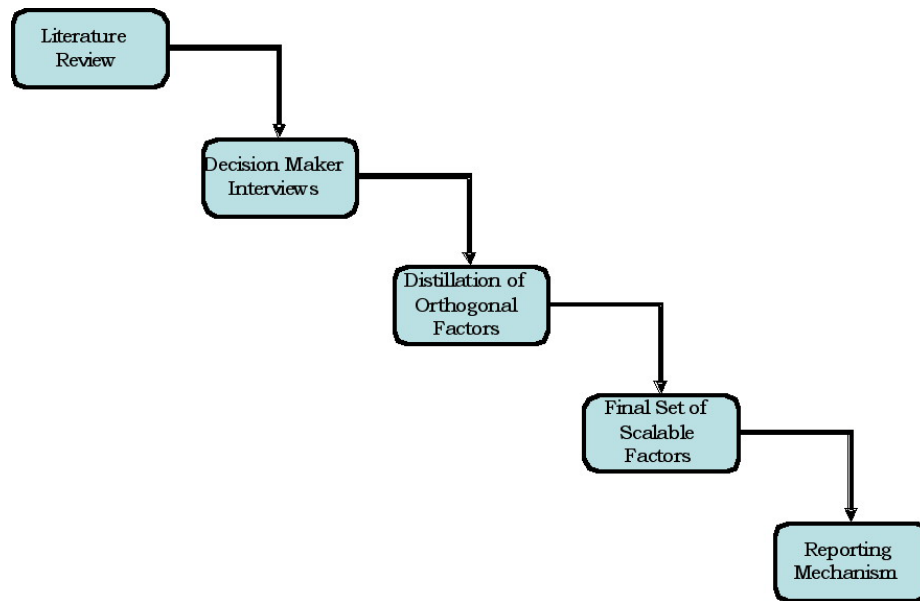



Figure 7.4-3. M&S Credibility Assessment Scale Development

7.4.1 Scale Workshop 1

The Scale Workshop 1 was held at LaRC on March 6 and 7. This was the only scale workshop with non-Topic Working Group members (aside from the facilitator) in attendance. The objectives were

1. Topic Working Group members understand the experiences (good, bad, and ugly) of other groups with M&S scales.
2. Topic Working Group refines the plans for the pilot to ensure optimal feedback on the scales.
3. Topic Working Group develops a plan to revise the scales.

All Topic Working Group members attended (except for Maria Babula and Kenneth Johnson, who joined the Topic Working Group later). Non-Topic Working Group members in attendance were Hal Bell (OCE), Steve Blattnig (LaRC), Charles Dunton (LaRC facilitator), Lawrence Green (LaRC), Scott Harmon (Zetetek & DMSO), Hans Mair (DoD), William Oberkamp (Sandia National Laboratories, Albuquerque), and Simone Youngblood (DMSO)

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 28 of 112

The first morning was devoted to presentations and discussions of various scales:

1. CFD Simulation Readiness Level (Kevin Tucker, MSFC) [ref. 14].
2. Predictive Capability Maturity Model (William Oberkampf, Sandia National Laboratories, Albuquerque) [ref. 13].
3. Validation Process Maturity Model (Scott Harmon, Zetetek) [ref. 12].
4. A2 scale (Steve Blattnig, LaRC) [ref 15].
5. A3 scale (Joe Hale, MSFC) [ref. 16].

A brainstorming session was held in the afternoon on identifying ideal characteristics of an M&S Scale. The next morning, the Topic Working Group identified the following information needed from the pilots in order to inform our decisions on revising the M&S Standard:

1. Need perspectives of both practitioners and decision-makers.
2. Is the M&S Standard being interpreted uniformly?
3. How well does the scale work for single M&S versus coupled M&S?
4. How much work and cost is this going to add?
5. How is credibility understood?
6. Can the M&S Standard be used to brief M&S work to management?


A plan for completing the scale was developed, with a target completion date of May 31.

The Topic Working Group formed three subteams to work specific issues raised at this workshop:

- Pilot Questionnaire: Zang
- M&S Credibility Literature: Mehta, Hale, Sylvester
- Information Quality Literature: Davis, Bertch, Mosier

These teams' products were developed and discussed during subsequent weekly Topic Working Group teleconferences.

The M&S Credibility Literature Subteam recommended the following papers be read by the full Topic Working Group: Mehta [ref. 17], Fogg & Tseng [ref. 18], Tseng & Fogg [ref. 19], and Balci [ref. 20]. The Information Quality Literature subteam did not recommend any papers to be read by the full Topic Working Group. However, the Wang & Strong [ref. 21] paper had been strongly recommended by both Scott Harmon and William Oberkampf at the workshop.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 29 of 112

This first workshop and the off-line work and weekly meetings over the following month or so constituted the literature review phase.

This first workshop led the Topic Working Group to conclude the following Finding:

F-7. There is a substantial literature on M&S credibility and/or scales, several other attempts, and numerous “lessons learned” on this subject.


7.4.2 Scale Workshop 2

The Scale Workshop 2 was held at JSC on April 24 and 25. The objectives were

1. Review the results of the Pilot Scale Questionnaire (Section 11.2).
2. Review the draft Pilot Questionnaire.
3. Make the major decisions on the credibility assessment scale, namely
 - a. What is the goal of the credibility assessment scale in the M&S Standard?
 - b. What would a candidate Concept of Operations (Use Case) be for the use of a credibility scale?
 - c. What are we trying to measure with the credibility assessment scale?
 - d. What are the key steps associated with M&S development that can be used for assessing credibility?
 - e. What key features/attributes of M&S can be used to assess credibility?
 - f. What model/architecture provides an efficient method for a credibility scale?

All Topic Working Group members attended (except for Kenneth Johnson).

The workshop contained two brainstorming sessions: one on possible architectures for the credibility assessment scale, and another on potential factors in the credibility assessment scale. (The Topic Working Group used the terms “architecture” and “model” interchangeably.) The results for the architecture brainstorming sessions are captured in [Figure 7.4-4](#). The focus of the discussion was whether the factors in the credibility assessment scale are considered as a set of serial (or sequential) products or processes, or whether the factors represent products produced or processes conducted in parallel. The implication of the serial model is that no factor is evaluated (or scored) unless all factors that precede it are evaluated at an acceptable level. With the parallel architecture, however, each factor is scored independently of the other factors. (The Permanent M&S Standard [ref. 5] uses the parallel architecture; an example of a scale with a serial architecture is given in ref. 22.)

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 30 of 112

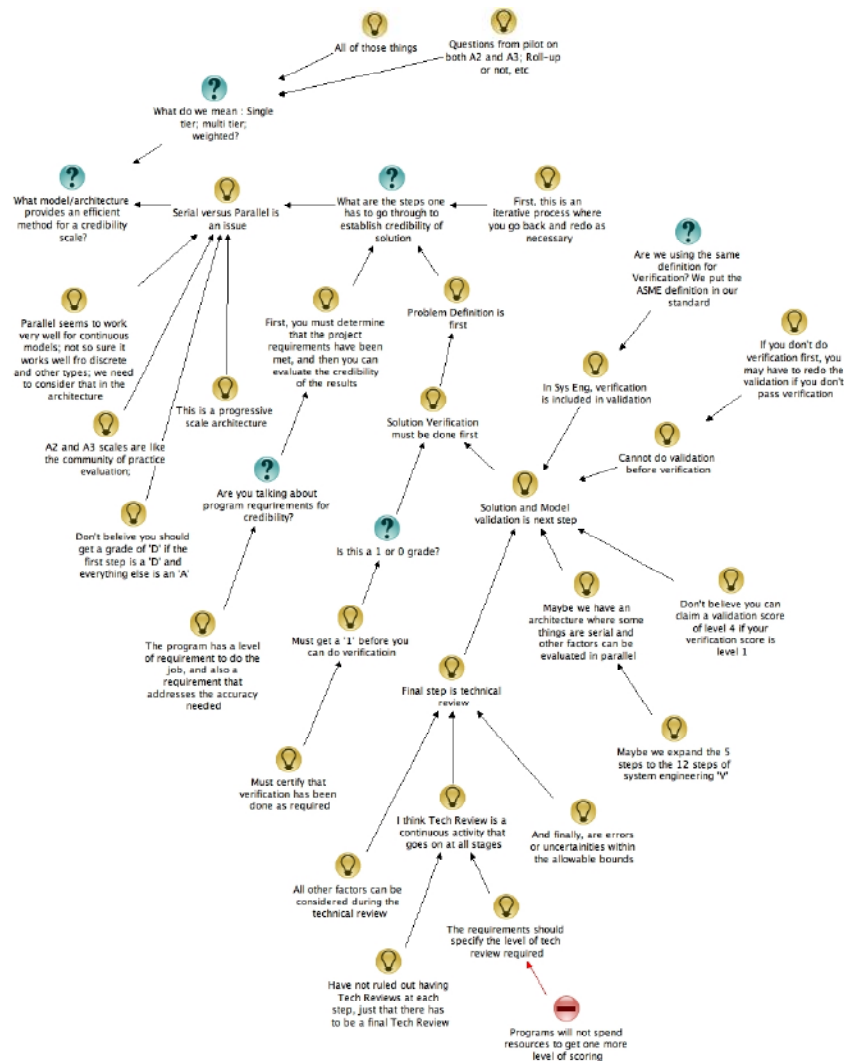



Figure 7.4-4. Brainstorming on the Best Architecture

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 31 of 112

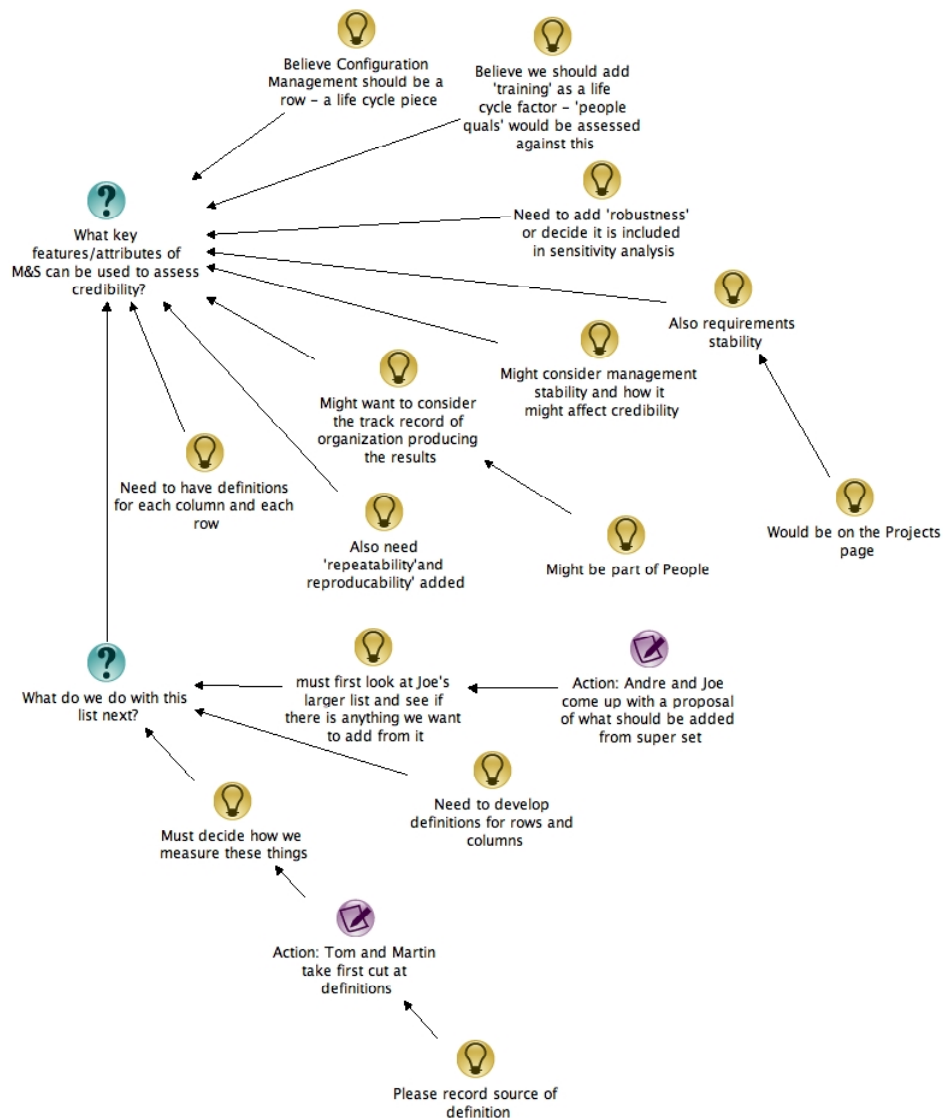



Figure 7.4-5. Brainstorming on the Key Factors for Assessing Credibility

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 32 of 112

Principal decisions that were made (by supermajority vote) were

1. The M&S Standard and the credibility assessment scale must cover all types of M&S.
2. The Topic Work Group should include some processes along with products as potential factors in the credibility assessment scale.

[Figure 7.4-5](#) illustrates the results of the brainstorming on key factors for assessing credibility. The list of candidate factors was over a hundred. The follow-up action was for the Topic Working Group to compile an organized list of all the candidate factors, with a crisp definition for each factor. This list ended up with 98 candidate factors.

Leading up to the subsequent workshop, the two important off-line activities for the Topic Working Group members were (1) to conduct and assess the implication of the interviews of decision makers, and (2) to rate the relevance of the candidate factors to assessing the credibility of the M&S results. The ratings for (2) were performed on the following scale:

1. Keep it (gotta have)
2. I'm on the fence (nice to have)
3. Get rid of it (not needed)


7.4.3 Scale Workshop 3

The Scale Workshop 3 was held at KSC on May 22 and 23. The objectives were:

1. Credibility Factors are identified at the 90–100 percent level.
2. The Hierarchy is identified at the 90 percent level.
3. Progress has been made on Level Definitions for the Credibility Factors.
4. Progress has been made on the roll-up algorithm for evaluation.

All Topic Working Group members attended¹. Much of the first day was occupied with reviewing the results of the decision-maker interviews (Section 11.3) and the candidate factor ratings. The full list of factors that were rated on the 1–3 scale and described at the end of Section 7.4.2 is given in [Figure 7.4-6](#). These factors were grouped into the 11 general areas highlighted in yellow in the lower right part of the figure. The factors highlighted in green were those that rated the highest, and those highlighted in orange were the “near misses.” Subsequent discussions concentrated on these 16 most highly rated factors. There followed some consolidation and refinement of the definitions of the factors. By the end of the workshop the

¹Jody Woods was present for only a few hours due to travel complications.


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 33 of 112

Topic Working Group succeeded in down-selecting the list to the 9 factors judged to be the most important. These were grouped into 3 categories.

The second objective focused on the details of the parallel architecture that had been decided at Workshop 2.

1 Robustness	4 Configuration Management	6 Technical Review
1 Sensitivity Analysis	4 Software Configuration Management	7 Accuracy
1 Margin (Margin Sensitivity)	4 Requirements Definition	7 Uncertainty Quantification
1 Risk	4 Project Management	7 Assertion Completeness
1 M&S Validity Under Load	4 Engineering	7 Accurateness of Model Response
1 Boundary Conditions	4 Support	7 Type A Evaluation of Uncertainty Using Statistical
1 Sensitivities	4 Process Management	7 Type B Evaluation Using Non-statistical Methods
1 Design of Experiments	4 Input Data (Data Management)	8 Input Validation
2 Fidelity	4 Execution Configuration (Config	8 Initial Conditions
2 Functionality	4 Output Data (Data Management)	8 Input Data Accuracy
2 M&S Verification (M&S V&V)	4 Has Model Been Maintained	9 People Qualifications
2 Adequacy of Representation	4 Infrastructure Set-Up	9 General Experience
2 Consistency of Representation	5 Requirements Stability	9 Education
2 Completeness of Representation	5 Reliability	9 Experience on the Specific M&S
2 M&S Causality	5 Reusability	9 Training for the Specific M&S
2 Interface Definition	5 Testability	10 Use Validation
2 Interface Functionality	5 Traceability	10 Range of Validity
2 Type of M&S	5 Usability	10 Use History
2 Software Control Flow	5 Remission of Coding Errors	10 Maturity
2 Data Variables Interdependency	5 Adherence to Coding Standards	10 Scale S Size of M&S
2 Input/Output Transformation Response	5 Adequacy of Execution Behavior	11 Repeatability
2 Software Interface Structure	5 Adequacy of Behavior During Fault/Failure	11 Reproducibility
2 Equations and Algorithms	5 Usability S User Interface	11 M&S Repeatability
2 Units of Measure	5 Supportability	11 Data Reduction
2 Standard Models Selection	5 Flexibility	11 Anomaly Resolution
2 Software Semantics	5 Maintainability	11 Data Analysis
2 Software Symbolic Accuracy	5 Extensibility	1 Sensitivity
2 Software Syntax	5 Portability	2 Verification
3 M&S Validation (M&S V&V)	5 Modularity	3 Validation
3 Output Validation	5 Specialty Engineering	4 Project Management
3 Predictive Capability	5 Operator Error Resolution	5 Process
3 Validity with Respect to Real-World Data	5 Software Error Resolution	6 Technical Review
3 Validity with Respect to Other M&S	5 Hardware Error Resolution	7 Accuracy
3 Predictive Validity	5 Data Management	8 Input Validation
3 Validity of M&S Subsystems (Structural	5 Complexity	9 Operator Qualification
	5 Report Generation	10 M&S Use
		11 Analysis

Figure 7.4-6. Initial Evaluation of Potential Credibility Assessment Scale Factors

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 34 of 112

The specific hierarchical architecture that was selected for the credibility assessment scale had 3 tiers:

1. Summary (top, or single, tier)
2. Categories (middle, or second, tier)
3. Factors (bottom, or third, tier)

The second tier—the categories—consisted of


1. M&S Development
2. M&S Operations
3. Supporting Evidence

(The category names identified here are those in the Revised M&S Standard. The preliminary names chosen at the Third Scale Workshop changed in the ensuing months, but their scope did not.)

The third tier—the factors—consisted of

1. Verification
2. Validation
3. Input Validation
4. Uncertainty Quantification
5. Sensitivity Analysis
6. Use History
7. Configuration Management
8. People Qualifications
9. Technical Review

After months of further discussion, the factors in the final credibility assessment scale that appears in the Permanent M&S Standard (compare with [Figure 8.0-1](#)) were remarkably similar to the preliminary version that emerged from the Scale Workshop 3. The major difference was that one of the factors—Technical Review—was later moved to a new, fourth tier, as a subfactor that influenced the assessments of the first 5 of the 8 remaining factors. A minor difference is that M&S Management was a narrower factor in the “KSC Scale”. There were also some name changes made subsequently, and Use History was moved from the M&S Development category to the Supporting Evidence category.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 35 of 112

The issues of whether and, if so, how to roll up the results was left for subsequent discussion. The main follow-up action was to provide level definitions for the factors. This was an important step because it would reveal whether or not a factor could actually be assessed objectively.

The results from the pilot studies—both the Pilot Questionnaires (Section 11.4) and the Pilot Reports (Section 7.3.4)—were reviewed during July.

7.4.4 Scale Workshop 4

The Scale Workshop 4 was held at GSFC on August 8 and 9. The objectives were

1. Finalize the credibility assessment scale to be used in the M&S Standard
 - Goal
 - Structure
 - Categories/factors/subfactors (Tiers 2-4)
 - Roll-up
 - Level Definitions
2. Complete substantial drafts of Sects 4.7 and Appendix A in the M&S Standard.


Most Topic Working Group members attended; Jody Woods and Kenneth Johnson were absent. Objective #1 was only partially completed, and Objective #2 was not addressed as Objective #1 needed to be completed first.

The goals identified for the credibility assessment scale were

- To define a common language by which credibility can be assessed,
- To inform decision-makers about the credibility of the current M&S results using the common language,

Underlying assumptions that were agreed upon were

- Credibility cannot be measured directly,
- Credibility assessment is accomplished by reviewing factors,
- The credibility assessment scale measures key factors that contribute to credibility,
- The factors are orthogonal, or nearly so,
- For factors that correspond to processes, the credibility assessment is based upon the quality of the process outputs,

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 36 of 112

The decision was made at this workshop to alter the structure of the credibility assessment scale by adding subfactors, but this was only for the purpose of moving Technical Review to a subfactor position for the first five factors.

Other major decisions made at this workshop were:


- Roll-up to a single number, but require it be accompanied by lower Tier information.
- Ignore Categories in any reporting.
- Do not require reporting of subfactors as a primary report, although they will necessarily be in backup.
- Use weighting in roll-up.
- Program defines weights, and M&S Standard Section 4.1 requires that the Program documents weights with rationale.

Some progress was made on the level definitions at the Scale Workshop 4. However, as the level definitions were refined, there were discussions about the distinctions between some of the factors. These discussions continued (as a secondary priority) into October, leading to significant changes in the focus of the factors. As the Topic Working Group learned, writing good level definitions is probably harder and more time-consuming than picking the factors themselves.

After this workshop, the top priority of the Topic Working Group was resolution of the comments from the formal Agency-wide review (Section 7.5).

7.4.5 Sandia National Laboratories-NASA Workshop

In between Scale Workshops 3 and 4, a 1-day meeting was held in Albuquerque, NM between several members of the Topic Working Group (Gary Mosier, Martin Steele, Andre Sylvester, Thomas Zang (via telephone)), Hal Bell of OCE, three individuals from Sandia National Laboratories (David Percy, Martin Pilch, William Oberkampf), and one person from Los Alamos National Laboratory (François Hemez). The purpose of the meeting was to compare experiences with scales. Sandia National Laboratories personnel applied the draft of the NASA credibility assessment scale to one of their applications (a legacy weapon exposed to a fire during an incident). The Topic Working Group discussed two applications of the PCMM to NASA M&S—the Orbiter RCC Impact Damage Threshold Assessment application (Section 7.2.2) and an integrated thermal-structural model for the James Webb Space Telescope. The Topic Working Group also gave Sandia National Laboratories personnel electronic copies of five other assessments of the PCMM to NASA applications, but there was only time at the meeting for discussion of the two mentioned above. The main conclusion that all attendees drew from the meeting was that clear terminology in the factor definitions and the level definitions was essential, and that this is a very difficult task. There were numerous specific comments and

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 37 of 112

questions from Sandia National Laboratories personnel on the details of the credibility assessment scale.

7.5 Formal Agency-wide Review

7.5.1 Interim M&S Standard Review

The formal NASA review of the Interim M&S Standard opened on May 11, 2007 and comments were due July 17, 2007. However, comments from several Centers were not received until the latter half of August 2007. A total of 377 comments were received in the form of a Comment Resolution Matrix.

The Topic Working Group's activities from mid-August through mid-November 2008 focused on the disposition of these comments. Most of this work was done via WebEx sessions. There was a face-to-face meeting at JPL on September 5–6, 2007. There also was a two-day teleconference on October 10–11, 2007.


Although the focus during this period was on the disposition of the formal comments, work continued on refinement of the level definitions in the credibility assessment scale.

The Topic Working Group either concurred (made exact change requested) or partially concurred (made a change similar to that requested) on 251 comments. The Topic Working Group concurred on 88 comments, and 37 comments did not request a substantive change (e.g., addressed style issues).

Comments were divided in major and minor comments. Subteams developed disposition recommendations for minor comments. The Topic Working Group reviewed these off-line, and these recommendations were adopted without discussion unless a Topic Working Group member objected to the recommendation. All major comments were discussed by the Topic Working Group. A supermajority was required in order to decide on the disposition of a comment. This entailed rather lengthy deliberations at times in order to reach a decision.

The disposition of each comment, along with the rationale for those comments with which the Topic Working Group did not concur, was recorded in the Comment Resolution Matrix. Each decision was conveyed to the commenter by the Topic Working Group member from the commenter's Center. Commenters objecting to the disposition of their comment(s) needed to persuade their Engineering Director to sustain their objection for this to be brought to the EMB for resolution.

[Table 7.5-1](#) records the votes of the Topic Working Group members on the Revised M&S Standard (Version 5). Because the credibility assessment scale was more controversial than the rest of the document, the vote was broken into two parts: (1) the scale and its related

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 38 of 112

requirements, and (2) all the rest of the M&S Standard. Three options were available (a) approve; (b) can live with it (but have some reservations); or (c) disapprove.

Table 7.5-5. Topic Working Group Votes on the Revised M&S Standard

Center	Representative	Vote on Credibility Assessment Scale	Vote on Remainder
ARC	Unmeel Mehta	Disapprove	Can Live With
GRC	Maria Babula	Approve	Approve
GSFC	Gary Mosier	Can Live With	Can Live With
JPL	William Bertch	Disapprove	Approve
JSC	Andre Sylvester	Approve	Approve
KSC	Martin Steele	Approve	Approve
LaRC	Lawrence Green	Approve	Approve
MSFC	Joe Hale	Can Live With	Approve
SSC	Jody Woods	Approve	Approve


The reasons for the two disapprove votes on the credibility assessment scale were as follows

ARC

Objective h [assure that the credibility of models and simulations meet the project requirements] is not met; there are questionable statements, such as “collectively they are nearly orthogonal, i.e., Independent factors;” the roll-up is arbitrary; the decision maker is asked to assess credibility; etc. See Section 10.2 for the alternative viewpoint on Version 5.

JPL

The input from the Chief Engineers from the various JPL directorates was that JPL did not want to use the credibility assessment scale. As noted in formal comment #71, two studies and two decision-maker interviews at JPL did not see a need for the credibility assessment scale. This is not to say that the credibility assessment scale is bad, it is just that JPL is a very "flat" organization and our decision makers are technical -- so they want to see the detailed numbers for verification and validation (V&V), robustness, etc -- not a mapping to the levels associated with the factors in the credibility assessment scale.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 39 of 112

The Topic Working Group submitted Version 5, along with the completed Comment Resolution Matrix, to the NASA Technical Standards Program Office on November 16, 2007 as their final deliverable. At this point the decision-making shifted to the EMB and the OCE.

7.5.2 Major Changes in the Revised M&S Standard

This subsection summarizes the major changes in the Revised M&S Standard (with respect to the Interim M&S Standard) as a result of the pilot studies, the workshops on the credibility assessment scale, and the NASA-wide review.

Section 1 Changes

There were numerous changes to increase the Scope understandability. (The changes have not changed the Scope.)


- A single goal and eight objectives replaced the previous two goals and five objectives.
- The objectives were drawn directly from Diaz Action #4 and the NASA Chief Engineer memo of September 1, 2006.
- The former Tables 1 and 2 were deleted.
- The relationship of this standard to NPR 7150.2 was clarified in Section 2.4.
- The applicability of the M&S Standard to COTS, GOTS, and MOTS tools was made explicit.

Section 3 Changes

- Numerous acronyms were added.
- Numerous definitions were added.
- Some of the existing definitions were modified.

Section 4 Changes

- Many requirements had language changes to increase clarity.
- Several new requirements were added, requiring documentation for some aspects of the new credibility assessment scale that were now covered by previous requirements (Reqs. 4.1.4, 4.1.6, 4.3.8, and 4.3.11).
- Related new requirements were 4.3.9 and 4.3.10.
- Previous Req. 4.1.1 was placed in the text as an expectation, per guidance from the TSPO.
- Previous Req. 4.1.7 was deleted.
- Previous Req. 4.1.5 was augmented and moved to Section 4.2.


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 40 of 112

- Previous Req. 4.4.5 was split into four separate requirements that covered the same material.
- All but one of the previous requirements (4.5.3) in Section 4.5 were dropped because of concerns about cost (i.e., eight requirements were dropped in Section 4.5).
- Previous Reqs. 4.6.3, 4.6.4, and 4.6.5 were dropped.
- Req. 4.6.3 was added.
- Previous Req. 4.7.1 was split into two requirements.
- Req. 4.7.3 was added.
- The Summary Credibility Scale level definitions were deleted.
- Previous Reqs. 4.8.1, 4.8.2, and 4.8.5 were consolidated into a single requirement, now Req. 4.8.1.
- Reqs. 4.8.2 and 4.8.3 were modified in accordance with changes to the credibility assessment scale.

Appendix A Changes

This was completely rewritten.

- There was no longer a Summary Credibility Scale.
- There was now just one credibility assessment scale, not two credibility assessment scales.
- The new credibility assessment scale had more factors than the A2 scale and fewer than the A3 scale.
- The new credibility assessment scale had some factors that were present in the previous credibility assessment scales.
- The new credibility assessment scale contained some factors that were not present in either previous credibility assessment scale.
- The new credibility assessment scale had five levels (versus four levels for the previous credibility assessment scales).
- The roll-up to a since number was a weighted average of the multidimensional factors, whereas previously it was the minimum score of the factors.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 41 of 112

Appendix B Changes

The Requirement Traceability Matrix was placed online. A Compliance Matrix was used instead in Appendix B.

7.6 Engineering Management Board Review

The EMB review occurred from January 22, 2008 through May 7, 2008. Initially, five Centers concurred, and five Centers non-concurred. The principal concerns were:

1. A guidebook was wanted and not a standard.
2. Did not want any type of scale.
3. Felt the current credibility assessment scale was too subjective.
4. Felt many requirements were unverifiable.

The major issues were discussed at the May 7, 2008 EMB meeting.

7.6.1 Major Changes in the Permanent M&S Standard

The following major changes in the Permanent M&S Standard (with respect to the Revised M&S Standard) were made as a result of this review:

Scope


- Added a new Appendix (now labeled Appendix A, with the previous Appendix A moving to B, and the previous Appendix B moving to C) describing a sample M&S risk matrix.
- Added sentences to Sections 1.2 and 4.1 referring to this new appendix.
- Inserted a new Req. 4.1.1 and modified Req. 4.1.2 (formerly 4.1.1) for consistency with the new approach to scope determination.

Role of Technical Authority

- The role of Technical Authority was strengthened (in Section 4.1) *with respect to*
 - Determining which M&S are in scope through the risk assessment,
 - Determining the level of detail appropriate for meeting the documentation requirements, and
 - Determining the objectives and requirements for the M&S products (Req. 4.1.3).

Documentation Detail

- The sentence “*The required documentation aspects for an activity that was not conducted may be simply satisfied by recording that the activity was not conducted.*” was removed from the end of Section 4.0. (Some had interpreted this sentence to mean that all

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 42 of 112

requirements for documentation could be satisfied merely by recording that the documentation was not done.)


- The sentence “Some requirements, in particular, 4.1.5, 4.2.6, 4.2.8, 4.3.6, 4.4.1, 4.4.2, 4.4.4, 4.4.5, 4.4.6, 4.4.7, 4.4.8, and 4.4.9, are to be interpreted as meaning that the activity in question is not required per se, but that whatever was done is to be documented, and if nothing was done a clear statement to that effect is to be documented.” was added, and wording changes were made to each of the listed requirements. This identifies requirements referring to documentation of activities, such as uncertainty quantification, that are not required; only the documentation of what, if anything, was done is required in these cases.

Credibility Assessment Scale

- The roll-up from the eight factor scores to the overall score was changed from a weighted roll-up to the minimum score across the eight factors. This resulted in the deletion of 2–3 pages of text and deletion of three references and acronyms.
- A paragraph was added to Section 4.7 to clarify the role of the Scale: “The operational concept of the credibility assessment scale is that the presentation of any results from M&S to a decision maker include (1) the best estimate of the results, (2) a statement on the uncertainty in the results, (3) the evaluation of the results on the credibility assessment scale, and (4) any explicit caveats that accompany the results. (An example of such a caveat would be use of the model in violation of its assumptions.) The decision maker then makes his/her own assessment of credibility based upon all four pieces of information in the context of the decision at hand. Just to emphasize this fundamental point, the credibility assessment scale does not purport to measure credibility; rather, it assesses the M&S results, and the rigor of the processes used to produce them, against key factors that affect the credibility judgment. The fundamental premise of this approach is that as a general rule, the more rigorous the key processes used for generating the M&S results, the greater the credibility of the M&S results, all else (including the estimated uncertainty) being equal.”
- Four clarifications were added to the text of Appendix B to reduce ambiguity in the level definitions.

Verifiability Clarifications

- Aegis, Inc. provided a detailed verifiability assessment of every requirement in the M&S Standard. As a result of their comments, minor wording changes were made to about a dozen requirements.
- In addition, the former Req. 4.4.4 was folded into Req. 4.4.1, and the former Req. 4.4.8 was folded into the former Req. 4.4.5 (now Req. 4.4.4), to reduce ambiguity.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 43 of 112

New Requirement

- Req. 4.2.13 was added by the Topic Working Group during the review of the language changes. The Topic Working Group felt that this information also needed to be in configuration management along with the related information in Req. 4.2.12.

7.6.2 Risk Assessment Matrix


At a March 27, 2008 meeting, the NASA Chief Engineer suggested to the Chair of the Topic Working Group that the determination of which M&S lie within scope of the M&S Standard be performed by an assessment of the risk incurred by the use of the M&S results. [Figure 7.6-7](#) illustrates the sample risk assessment matrix that was added to the Permanent M&S Standard for this purpose. The standard leaves the choice of the number of levels for Decision Consequence up to the Program/Project. The sample uses Decision Consequence level definitions adapted from those in NPR 8000.4, which is the only specific risk matrix in a NASA guidance document. See Appendix D for the precise text in that NPR at the time this adaptation was made.

M&S Results Influence	5: Controlling	(G)	(Y)	(R)	(R)
	4: Significant	(G)	(Y)	(R)	(R)
	3: Moderate	(G)	(Y)	(Y)	(R)
	2: Minor	(G)	(G)	(Y)	(Y)
	1: Negligible	(G)	(G)	(G)	(G)
		IV: Negligible	III: Marginal	II: Critical	I: Catastrophic
		Decision Consequence			

Figure 7.6-7. Sample M&S Risk Assessment Matrix

7.6.3 Issuance of Permanent M&S Standard


The Permanent M&S Standard, incorporating the changes described in Section 7.6.1, was reviewed once more by the EMB in late June 2008. No objections were recorded. The NASA Chief Engineer formally issued the Permanent M&S Standard on July 11, 2008.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 44 of 112

7.7 Outreach

Over the course of this activity numerous presentations were made at conferences and professional society meetings on aspects of the M&S Standard. These include presentations at the

- Finite Element Modeling Continuous Improvement Workshop [Zang, October 2006]
- AIAA Standards Technical Committee [Mehta, January 2007]
- Simulation Interoperability Standards Organization (SISO) Simulation Interoperability Workshop [Zang, March 2007]
- Finite Elements in Fluids Conference [Green, March 2007]
- ASME Committee on Verification and Validation in Computational Solid Mechanics [Zang, March 2007]
- Joint Army, Navy, NASA, Air Force (JANNAF) Modeling and Simulation Subcommittee Meeting [Hale, May 2007]
- JANNAF Modeling and Simulation Subcommittee Meeting [Mehta, May 2007]
- Society for Modeling and Simulation International (SCS) Summer Computer Simulation Conference [Steele, July 2007]
- Sandia National Laboratories Workshop on Mathematical Methods in V&V [Zang, August 2007]
- SISO Simulation Interoperability Workshop—VV&A Summit [Hale, September 2007]
- SISO Simulation Interoperability Workshop—VV&A Summit [Zang, September 2007]
- AIAA Nondeterministic Approaches Conference [Luckring, April 2008]
- SISO Simulation Interoperability Workshop [Bertch, April 2008]
- SISO Simulation Interoperability Workshop [Hale, April 2008]
- SISO-SCS International Simulation Multi-conference [Steele, June 2008]

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 45 of 112

8.0 Overview of the Permanent M&S Standard

The overall goal of this Standard is to ensure that the credibility of the results from M&S is properly conveyed to those making critical decisions. That is, requirements are identified to ensure the development, operation, and documentation are properly addressed, but the critical requirements specify what is to be presented to the decision-makers. Having requirements on the *presentation* is a rather unique approach for a standard, but gets to the heart of the issues raised in the CAIB report where the information was not properly conveyed to the decision-makers.

The main body of the M&S Standard consists of two parts as shown in [Figure 8.0-1](#). The first part addresses the conventional set of requirements for M&S projects. The second part addresses the use of a credibility assessment scale that was included to make the M&S credibility more apparent to the decision-maker with the anticipation that this can expose the risk associated with M&S-based decisions. This section discusses each part in turn, after discussion the Scope and Definitions sections.

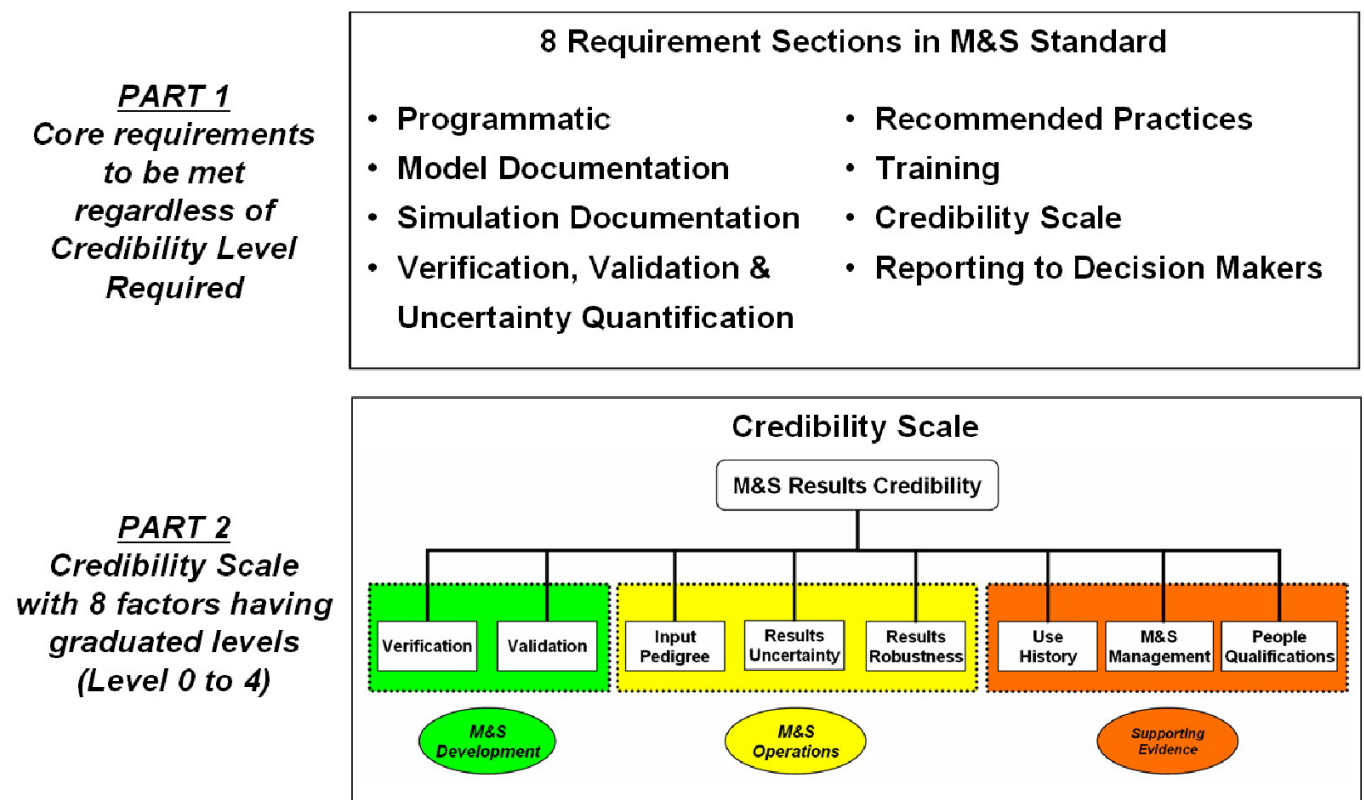



Figure 8.0-1. Two Parts of the M&S Standard

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 46 of 112

8.1 Scope

Determination and articulation of the scope of this standard was a challenging and iterative task. One month after the start of Phase 1, on June 14, 2005, the Development Team met with the OCE liaison at the time (Michael Blythe) to seek guidance on the scope. His general guidance was to stick to the intent of the Diaz Action #4 (i.e., on M&S used for decisions affecting human safety and mission success). The implication was that the Development Team should not take Diaz Action #4 as a license to address broader M&S concerns.


The major question that the Development Team had was whether the scope included software used for control systems and displays. This was ruled out of scope by the OCE liaison. The following information was used to facilitate the refinement of the scope during this June 2005 discussion:

M&S Uses in Engineering

- Technology Investment (*out of scope*)
Identify and evaluate candidate advanced technologies for future missions and systems
- Acquisition (*out of scope*)
Specify and acquire new systems
- Analysis & Design (*in scope, Priority Level 2*)
Evaluate and explore solution spaces for current and future systems and subsystems
- Test & Evaluation (*in scope, Priority Level 2*)
Evaluate/verify hardware & software artifacts
- Training (*probably out of scope*)
Produce learning in a user or participant
- Engineering/Operations (*in scope, Priority Level 1*)
Evaluate status/anomalies/corrective actions in operational systems

M&S Uses in Science

- Scientific Data Analysis (*out of scope*)
Process data from scientific instruments
- Scientific Understanding (*out of scope*)
Simulation of natural phenomena used for advancement of scientific knowledge
- Natural Phenomena Prediction (*partly in scope, Priority Level 3*)

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 47 of 112

Simulation of natural phenomena used for operational decisions affecting safety and mission success

Predictions of the operational environment that have a direct impact on the safety of personnel and NASA assets are in scope provided that NASA has primary responsibility for these predictions —space weather is an example


Simulations for which other agencies have primary responsibility are out of scope — Earth weather is an example

The guidance from Michael Blythe is recorded in the italicized comments in parentheses. A day after this meeting, Michael Blythe met with the Deputy Chief Engineer, Gregory Robinson, who confirmed this guidance, but did indicate that the stimulation environment (often consisting of M&S) for control systems was in scope.

Although the previously identified information provided general clarification to the sense of the scope, and appeared in a revised form in the Interim M&S Standard, they proved to be unclear to a number of the commenters during the formal review. As expected, there were numerous formal comments on the Scope and Applicability of the M&S Standard. Many of these comments fell into one of three categories: (1) objections to the breadth of the scope, (2) inability to understand the scope, and (3) questions about the relationship to NPR 7150.2, Software Engineering Procedures.

In responding to (1), the Topic Working Group sought additional guidance from Hal Bell, as this is a policy rather than a technical decision. Because of (2), the Topic Working Group chose to greatly simplify the articulation of the scope in the Revised M&S Standard; see Section 6 in the standard for the final language. As noted in Section 7.6.2, an additional refinement to the Scope—the connection to an M&S risk assessment—was made as a result of the EMB review.

For (3), the Topic Working Group determined that there was very little overlap, and no outright contraction, between the two documents. (An independent assessment conducted by Milton Lavin at JPL, albeit in the context of comparing the M&S Standard with JPL’s internal implementation guidelines for NPR 7150.2, reached the same conclusion.) In particular, the NPR has one general requirement to “test, validate, and certify software models, simulations, and analysis tools [requirement SWE-070]”, and it does not mention uncertainty quantification. On the other hand, the M&S Standard has a limited number of software-specific requirements such as providing version control and use of a configuration management system. Thus, the two documents are complementary, with the M&S Standard providing requirements for all the aspects of M&S that have more to do with the scientific method than with software engineering. Discussions with the NASA official responsible for NPR 7150.2 (John Kelly) led to inclusion of the following language in the M&S Standard: “implementation plans for NPR 7150.2 ... should ... address such M&S-specific issues as numerical accuracy, uncertainty analysis, sensitivity

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 48 of 112

analysis, M&S verification and M&S validation” to emphasize these M&S specific requirements.

As discussed in Section 7.6.2, the Permanent M&S Standard added the use of a risk assessment process for determination of which M&S are in Scope.

8.2 Definitions

The approach taken to the definitions that appear in Section 3.2 of the M&S Standard was to use definitions from the M&S community (and not from the systems engineering, software engineering, or statistics communities). Where available and appropriate, these definitions were extracted from Agency level directives as found on the NASA Online Directives Information System. In most other cases, they were taken or adapted from consensus publications, such as professional society guides.


8.3 Requirements

The requirements section consists of forty-nine requirements separated into eight subsections. The first six subsections provide the underlying activities that support the credibility assessment requirements in subsection 7, and subsection 8 addresses the reporting of M&S results to the decision makers.

The introductory material for each requirements section includes a discussion of the intent of the requirements in that section. Thirty-three of these requirements start with the words “shall document.” Twelve of these, in particular, 4.1.5, 4.2.6, 4.2.8, 4.3.6, 4.4.1, 4.4.2, 4.4.4, 4.4.5, 4.4.6, 4.4.7, 4.4.8, and 4.4.9, are to be interpreted as meaning that the activity in question is not required per se, but that whatever was done is to be documented, and if nothing was done a clear statement to that effect is to be documented.

The first requirements subsection addresses *programmatic activities*. The most fundamental activity is for the project management in collaboration with the Technical Authority to identify and document the critical decisions to be addressed with M&S and to determine which M&S are in scope. The latter determination should be based upon the risk posed by the anticipated use of the M&S, using the risk assessment approach discussed in Section 7.6.2. These requirements oblige the Project to: 1) identify the M&S that are in scope, 2) define the objectives and requirements for the M&S, and 3) develop a plan for the acquisition, development, operation, maintenance, and/or retirement of the M&S.

The second requirements subsection addresses the *requirements imposed on the model*, where model refers to the conceptual model, the mathematical model, and the computational model. The majority of these requirements address documentation for the assumptions, basic structure, mathematics, data sets, limits of operation, guidance in the proper use of the model, parameter

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 49 of 112

calibrations, model updates, and methods for uncertainty quantification for any data used to develop the model or incorporated into the model.

The third requirements subsection addresses the *requirements imposed on the simulation*. This includes requirements addressing the limits of operation, pedigree of the input data, processes for executing the simulations, processes for conducting analyses, assessment of the appropriateness of the simulation relative to its intended use, and use history of the M&S.

The fourth requirements subsection addresses the *verification, validation, and uncertainty quantification*. M&S practitioners typically understand the nuances of these requirements for their particular type of M&S. Specific emphasis is given to communicating the domains of verification and validation of the model to assure appropriate application of the model. Furthermore, documentation of the uncertainties in the results and their sensitivities is required.


The fifth requirements subsection of the M&S Standard addresses the use of *recommended practices*. The sixth requirements subsection addresses *training for developers, operators, and analysts*. (Both topics were explicitly specified in the Diaz Action #4.)

The seventh requirements subsection addresses the *credibility assessment scale*. The requirements specify that the M&S results and processes be assessed on the credibility assessment scale defined in Appendix B of the M&S Standard.

The eighth and final requirements subsection addresses the *reporting of results* to decision makers. This is the key activity driven by this M&S Standard. This is discussed in more detail in Section 8.4.

8.4 Credibility Assessment Scale

The operational concept of the credibility assessment scale is that the presentation of any results from M&S to a decision-maker would include: (1) the best estimate of the results, (2) a statement on the uncertainty in the results, (3) the evaluation of the results on the credibility assessment scale, and (4) any explicit caveats that accompany the results. (An example of such a caveat would be use of the model in violation of its assumptions.) The decision-maker then makes his assessment of credibility based upon all four pieces of information in the context of the decision at hand. Just to emphasize this fundamental point, the credibility assessment scale does not purport to measure credibility; rather, it assesses the M&S results, and the processes used to produce them, against key factors that affect the credibility judgment. The Topic Working Group stresses that the goal of this scale is to assist in the assessment of the credibility of the particular results at hand and not to assist in a broad certification (or accreditation as some prefer to call it) decision for a class of uses of the M&S.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 50 of 112

See Sections 7.2.3 and 7.3.2 for the background and basic decisions made the Development Team and the Topic Working Group on the two credibility assessment scales that appeared in the Interim M&S Standard. Section 7.5 provides a lengthy summary of the Topic Working Group's revision during Phase 2. The basic features of the credibility assessment scale are provided in the bottom half of [Figure 8.0-1](#). See the Permanent M&S Standard itself [ref. 5] for the detailed explanation of the factors, the level definitions and the roll-up process.

As noted in Section 7.4.3, the Topic Working Group considered nearly 100 separate factors, obviously far too long a list. It is well beyond the Magical Number Seven (Plus or Minus Two) rule of Miller [ref. 23]. Of all the many candidate factors that did not make the final list, two of those merit special comment—accuracy and fidelity. These candidate factors were suggested in some Decision-maker Interviews and Pilot Scale Questionnaire responses.

Accuracy rated in the top 16 at the KSC workshop. However, the Topic Working Group judged that it and uncertainty were not sufficiently orthogonal. Fidelity, on the other hand, did not rank in the top quartile.

8.5 Traceability

Section 6 lists the objectives for the M&S Standard, as provided by Diaz Action #4 and the memo from the NASA Chief Engineer. A traceability matrix that links the requirements in the M&S Standard with the Diaz Action #4 details the Chief Engineer's direction to include a credibility assessment scale is furnished in Appendix E.


Some aspects of Diaz Action #4 are covered rather sparsely by explicit requirements in the M&S Standard. These are to

- Identify best practices to ensure that knowledge of operations is captured in the user interfaces (e.g. users are not able to enter parameters that are out of bounds)
- Develop a process for user feedback when results appear unrealistic or defy explanation

It is a practical impossibility to construct a general method for ensuring that M&S are not used with parameters that are out of bounds (i.e., outside the limits of operation of the M&S). There are M&S Standard requirements (Reqs. 4.2.5, 4.2.7, 4.3.1, 4.5.1 and 4.8.1) and recommendations (4.3b, 4.3c, 4.5j, 4.6a) that establish precautions to reduce the likelihood of using M&S with parameters that are out of bounds. The items left to recommendations are ones that may not be technically or practically feasible in all cases.

Furthermore, the certification/recertification aspect of the following part of the Diaz Action #4 has deliberately not been addressed by means of an explicit requirement for certification.

- Develop process for tool verification and validation, certification, reverification, revalidation, and recertification based on operational data and trending

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 51 of 112

Rather, Req. 4.1.3a addresses this obliquely, leaving the decision on whether any type of endorsement, including certification, up to the Program/Project and the Technical Authority. This was done at the explicit direction of the OCE.

9.0 Observations, Findings and Recommendations


The Observations, Findings and Recommendations are listed separately for the Development Team and the Topic Working Group. The listed Recommendations are directed towards the NASA Chief Engineer unless otherwise identified.

9.1 Summary of Observations and Findings

The following Observations and Findings were identified earlier in this document, and are recorded below for completeness.

Development Team Observations and Findings:

- O-1. Development of a rigor scale is extremely difficult, even for a small, homogeneous group, and even restricted to M&S using PDE-based mathematical models.**
- O-2. The credibility assessment scale is not a standalone assessment of factors influencing credibility, but rather the credibility assessment scale plus the uncertainty statement combine to influence the credibility assessment by the decision-maker.**
- O-3. Frequent face-to-face meetings were essential for the initial formulation of the M&S Standard.**
- O-4. Many engineers and program managers at NASA are unaware of the intended hierarchy of the agency guidance documents.**
- O-5. Many engineers at NASA are unaware of standards that are relevant to their work.**
- F-1. Current NASA guidance is oriented towards control systems and displays. Quality assurance and configuration management are very well covered, but the unique, critical aspects of M&S are not addressed, for example, validation against experimental or flight data, and uncertainty quantification.**
- F-2. No federal agency has an M&S Standard, although the DoD has extensive M&S guidance, and the Nuclear Regulatory Commission has standards for control systems and displays.**
- F-3. Relevant M&S guidance is strongly focused on the development phase of the M&S life-cycle, and especially upon verification and validation. There is little guidance on the operations of M&S and virtually no guidance on the maintenance of M&S.**
- F-4. NASA has no policy nor any procedural requirements for M&S except for the software engineering aspects of M&S covered by NPD 2820.1B and NPR 7150.2.**

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 52 of 112

F-5. There does not presently exist an M&S scale with the specific objectives desired for the M&S Standard.

Topic Working Group Observations and Findings:

O-6. The credibility assessment scale is outside the formal Diaz Action #4.

F-6. The definition of a Topic Working Group consensus in the NASA Interim Directive on Interim NASA Technical Standards is not sufficiently precise.

F-7. There is a substantial literature on M&S credibility and/or scales, several other attempts, and numerous “lessons learned” on this subject.

9.2 Development Team Recommendations

R-1. NASA should integrate the M&S Standard into the NASA guidance hierarchy.

The initial review of existing M&S guidance and standards (Section 7.2.1) made it apparent that the M&S Standard was not tied to any existing NPD or NPR. The most logical existing NPR that could link to the M&S Standard would be NPR 7123.1A (NASA Systems Engineering Processes and Requirements). Either the M&S Standard should be linked to an NPD, a future version of NPR 7123.1A or to a forthcoming NPR on NASA standards.


R-2. NASA should coordinate with other organizations and professional societies to further mature the M&S Standard.

The development and operation of M&S, the analysis and presentation of M&S results, the proper training of M&S practitioners, the identification of recommended practices, and the need for assessing and conveying the credibility of M&S results to decision makers are not unique to NASA. These aspects of the M&S process are common to many other organizations. NASA should participate in activities directed towards standards that serve a broader M&S community.

R-3. NASA should sponsor development of Recommended Practices Guides.

While some M&S have well established and documented procedures, many others do not. Furthermore, existing guidelines may not cover new applications of the M&S. For example, models often require calibration, or numerical parameters need to be tuned for new problems. Knowledge of these procedures, calibrations, and tunings often resides in a small subset of workers. NASA should identify M&S domains that need Recommended Practice Guides and coordinate with professional societies, academia, commercial and international partners to develop them. (Domains may be organized according to type of M&S, by discipline, or by application.)

R-4. NASA should re-assess the requirements on recommended practices that were removed from the Interim M&S Standard.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 53 of 112

The Interim M&S Standard contained nine requirements for the Recommended Practices section. Eight of these requirements were deleted in the final Standard. What is left is merely the requirement to identify existing applicable RPGs. Greater discipline in the use of M&S is best fostered by development of new RPGs where needed. This can be a key element in training of developers, operators and analysts.

R-5. NASA should refine how submodels are treated in the credibility assessment scale.

The present version of the M&S Standard makes no distinction between individual models and integrated models consisting of multiple submodels. The roll-up of assessments of the individual submodels into the assessment of the integrated model is primarily an issue for the credibility assessment scale. The credibility assessment should eventually be refined to account for the additional issues associated with integration of submodels.

R-6. Information regarding credibility assessment scale usage should be collected to determine effectiveness and provide data for further revision.

In general, scales measuring the rigor, credibility, or similar aspects of M&S results have not received much use, and there is no consensus on such assessments. In particular, the credibility assessment scale in the M&S Standard has not been used. The immaturity of this particular field necessitates close monitoring of the impact of credibility assessment scale usage by NASA programs and the use of that information to update the credibility assessment scale. This is not a criticism of the present credibility assessment scale, but merely an acknowledgment of the state of such assessments; operational use is essential to advance the state-of-the-art.

R-7. NASA should clarify the operational meaning of the term “consensus” in NASA Interim Directive on Interim NASA Technical Standards.


As noted above in F-6 the directive requires a Topic Working Group “consensus”, but does not give a sufficiently precise definition of this term. A definition with a clear, operational meaning is needed.

9.3 Topic Working Group Recommendations

Some recommendations below are similar to the Development Team, but are recorded as well for reinforcement and/or clarification

R-8. NASA should sponsor the development of Recommended Practices Guides along disciplinary lines. This responsibility might best be delegated to the NASA Technical Fellows.

See R-3 for the rationale. The Topic Working Group supplemented this with a particular suggestion of who might be tasked with this responsibility. This does not necessarily mean that the NASA Technical Fellows should personally develop the guides appropriate for their

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 54 of 112

disciplines, but merely they should have the responsibility (and budget) for ensuring that they are produced.

R-9. NASA should collect data on the scope decisions, the cost impact and the credibility assessment scale usage of the M&S Standard.

This is a more general recommendation than R-6. The extension to collection of data on the scope decisions and cost impact was motivated by the large number of comments on these topics submitted as part of the Agency-wide review.

R-10. NASA should develop, by application domain, an M&S “validation lessons learned” database.


This information would be used to develop guidelines allowing designers to intelligently balance risk versus conservatism during program/project formulation. Solid data and rationale for design margins exist, in the form of written guidelines at the agency level, for only a few of the many application domains (i.e., disciplines). Of particular interest is the knowledge of why and by how much M&S results were in error before the models were tuned/correlated. NASA should also implement a process by which the guidelines are continuously re-evaluated and updated as the database grows.

R-11. An NPD and/or NPR should call out the M&S Standard.

See R-1 for the rationale.

R-12. Centers should share with each other their plans and other guidance for implementation of the M&S Standard.

The M&S Standard is a first-of-its-kind document, and there are few existing Recommended Practices Guides for M&S. NASA would make more effective use of the M&S Standard by sharing the individual Center implementation plans and guidance documents than by having each Center work this independently.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 55 of 112

10.0 Alternate Viewpoints

10.1 Alternate Viewpoint (Unmeel Mehta) on the Interim M&S Standard

See Appendix C for the minority opinion on the Interim M&S Standard.

10.2 Alternate Viewpoint (Unmeel Mehta) on the Revised M&S Standard


Because of the following four reasons, Version 5 is not appropriate as a standard.

1. The method to assess the credibility of M&S results presented in credibility assessment scale (Appendix A) and the associated requirements in Section 4 are questionable. The method is subjective, complex, and unsound. It leads to non-uniform/non-standard credibility/quality assessment of M&S results. The method does not provide the credibility assessment, instead the decision maker is asked to make the assessment. The method and requirements do not fulfill Objectives (g) and (h).
2. Version 5 does not focus on outcome. The method for credibility assessment is output based rather than performance (outcome) based. M&S results are procured for engineering efforts, including for making critical decisions. The NASA policy is to prefer use of performance (outcome-based) standards in procurement activities over design or process (method-based) standards (NPD 8070.6B for Technical Standards).
3. The development of processes—a stated objective—is not met. The required processes for validation, verification, uncertainty quantification, certification, etc. are not presented. Among a total of 49 requirements, 39 requirements are for documentation. Without proper processes, documentation by itself does assure quality of M&S results.
4. Version 5 does not meet the definition of a Standard. It does not address all stated objectives. Waivers, tailoring, factor weights, and the availability of option not to quantify uncertainties make this version a non-standard. A uniform engineering and technical requirement, a necessary condition for a standard, is not established. Please see the definition of NASA engineering standard and the applicability statement in the Applicability section of SAE AS 9100, Rev. B.

The questionable method for credibility assessment of M&S results for critical decision, the failure to address Objective (h), the failure to focus on performance, non-uniform applicability of Version 5, and the focus on documentation provide questionable value to the program and the Agency.

10.3 Alternate Recommendation (Unmeel Mehta)

The composition of the Topic Working Group and the selection of members for this Group should be done very judiciously to achieve a highly successful outcome.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 56 of 112

The composition of Topic Working Group and the selection of members for this Group should be done judiciously, including built-in checks-and-balances, to achieve a highly successful outcome. The guidance for selection of Topic Working Group is as follows: The Topic Working Group to develop a NASA M&S Standard should have seven members with extensive experience in development and use of M&S and three members with extensive experience as program or project leads who have used M&S results for engineering and for critical decisions. Among the six M&S experts, three disciplines should be addressed, with two experts in each of the three disciplines. Computational Fluid Dynamics (CFD) and Computational Solid Mechanics (CSM) are examples of disciplines to include. The Topic Working Group lead, the tenth person, should be the person with M&S expertise and from the discipline that is the most used within the Agency. This person should also not be the leader of the Development Team for M&S standard and should be from a center other than that of the Development Team lead. The latter person should be a Topic Working Group member. There should also be only one representative from each Center, with each member having the voting right.

11.0 Other Deliverables


This section first lists the final documents delivered under this task. Then, since the pilot studies in Spring 2007 were such important input to this process, the two questionnaires are recorded and the results from these are summarized. The decision-maker interviews conducted during the same period were also influential. These are also recorded here. Some of the background information on the questionnaire that is redundant for this report has been omitted below; bracketed, italicized notes indicate the locations and content of these deletions.

11.1 Standard Documents

The deliverable from the Development Team for their part of this task was Version 3, which was submitted to OCE on August 15, 2006.

The primary deliverable from the Topic Working Group for their part of this task was the document—the Revised M&S Standard—for the proposed NASA Standard for Models and Simulations (which underwent some subsequent modifications as a result of the EMB review). This document was delivered to the NASA Technical Standards Program Office on November 16, 2007, along with the Comment Resolution Matrix, which documented the Topic Working Group decisions on the formal comments submitted during the NASA-wide review.

The secondary deliverable from the Topic Working Group was the document for the Interim NASA Standard for Models and Simulations, which was delivered to the NASA Technical Standards Program Office on October 11, 2006.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 57 of 112

After implementing the changes resulting from the EMB review, the final document—the Permanent M&S Standard—was delivered to the NASA Technical Standards Program Office on May 23, 2008. Also included were the updated Comment Resolution Matrix, the Traceability Matrix (for posting on the web), an overview briefing for M&S practitioners, and an assessment questionnaire for evaluating the M&S Standard in actual use on M&S projects (accompanied by a worksheet for assessing its cost impact of the M&S Standard).

The Traceability Matrix is included in this report in Appendix E. The assessment questionnaire was adapted from the ones used in the pilot studies of the Interim M&S Standard (given here in Appendices E and G). The cost impact worksheet requires that one of the following three assessments be made for each requirement in the M&S Standard:


- Do This Already
- Don't Do This: Minimal New Cost
- Don't Do This: New Cost Driver

11.2 Pilot Scale Questionnaire

The questionnaire is provided in Appendix F and summarizes the results and conclusions. In interpreting the results, note that term “category” used in the questionnaire is equivalent to the term “factor” used in the final credibility assessment scale. (Compare the guidance in the second paragraph of the Introduction segment of Section 11.2.1 with the terminology used in the Interim M&S Standard.)

Two of the responses to this questionnaire were from members of the Development Team responsible for the A2 scale. Their responses are not included in this summary. There were 14 other responses. The answers to the multiple-choice questions are summarized in two different graphical formats.

[Figure 11.2-1](#) uses a bar plot to summarize the data, whereas [Figure 11.2-2](#) uses a mosaic plot. For example, on question #5, the responses were: 7 for (a); 5 for (b); and 3 for (c). The questions not shown in this figure were open ended rather than multiple choice. Note that the responses on this questionnaire are only indirectly relevant to the credibility assessment scale in the Permanent M&S Standard, as that credibility assessment scale is substantially different from both credibility assessment scales in the Interim M&S Standard.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 58 of 112

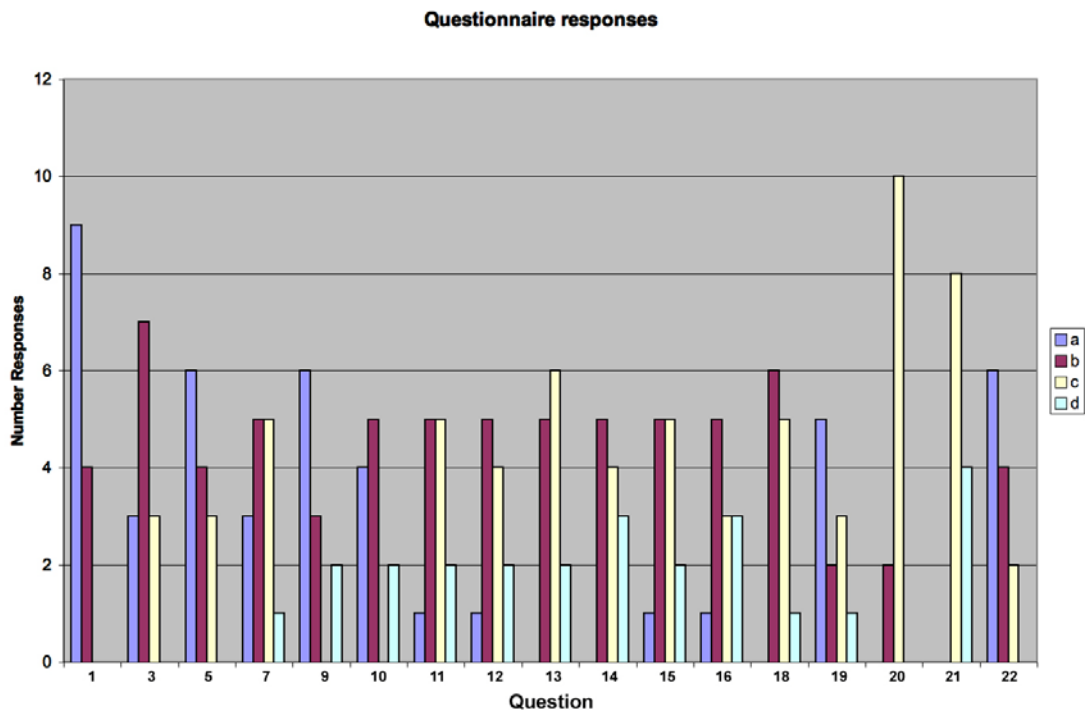


Figure 11.2-1. Bar Chart of Results from the Pilot Scale Questionnaire

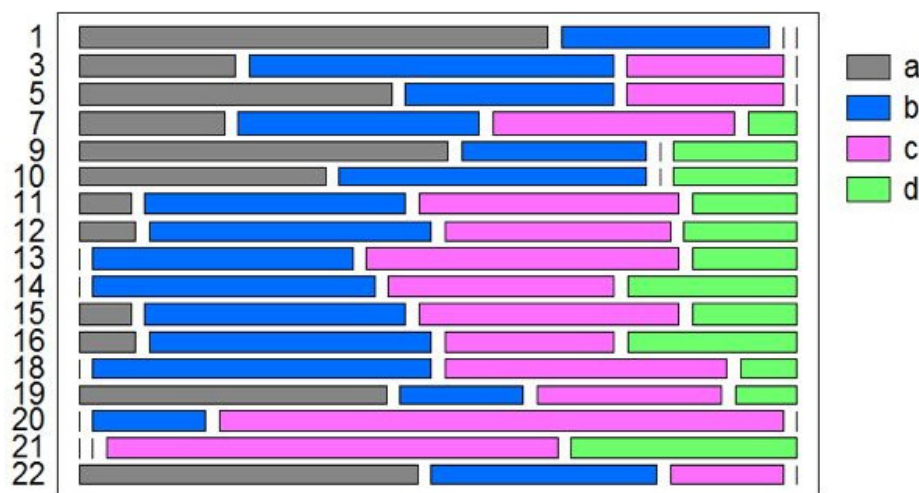



Figure 11.2-2. Mosaic Plot of Results from the Pilot Scale Questionnaire

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 59 of 112

Individual Topic Working Group members voiced their conclusions below from the review of the Pilot Scale Questionnaire:

- The current scales are not good enough based on the results of Q11–Q16
- Number of categories should be less than 10 (Q8)
- No one wants to give equal weight to all categories (Q20)
- No one wants to give the minimum —should be weighted, or mean and min/max (Q21)
- A2 is more understandable than A3 (Q1 and Q2)
- Both scales work about as well for coupled, with a slight advantage to A21 (Q9 and Q10)
- Split opinion on question of roll-up and whether it meets the goals (Q11–Q16)
- There were some comments on what were the most important categories that we should pay attention to

Note that these were individual conclusions and not necessarily consensus Topic Working Group conclusions.

11.3 Decision-Maker Interviews


The Decision-Maker Interview Guide itself is provided in Appendix G. At the Scale Workshop 3, the Topic Working Group members reported the main points from the interviews that they had conducted. Then, the following list was made of potential factors that were emphasized in the decision-maker interviews:

- Accuracy–error bars–uncertainty
- Fidelity of model
- Verification and validation
- Qualifications of people doing the work
- Fit intended use
- Validation of input
- Independent review or analysis
- Validate against real-world data
- Traceability to past knowledge
- Use history of model

11.4 Pilot Questionnaire

The principal product from the Pilot Studies was the responses on the Pilot Questionnaire. The questionnaire itself is provided in Appendix H.

A constant refrain during the years of development of this standard was that it would cost too much. Hence, the purpose of questions 7–11 and 13–16 was to elicit estimates of the impact of this standard on the cost of the M&S. There were 8 responses to questions 7–11. All used the A2

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 60 of 112

scale as the context for their estimates. [Table 11.4-1](#) gives the quantitative estimates collected from the pilot teams for the estimated cost impact to achieve Level 1 on the A2 scale.

Table 11.4-1. Raw Data for Cost Estimates for Level 1

	Minimum	Most Likely	Maximum
ARC-Holst	1.05	1.08	1.10
GSFC-Liu	1.05	1.10	1.20
JPL-Aquarius	2.00	2.20	3.00
JPL-Kepler	1.00	1.00	1.00
JPL-MER-EDL	1.05	1.06	1.10
JPL-MER-Thermal	1.00	1.00	1.00
KSC-SpaceNet	1.40	1.50	1.60
MSFC-Nix	1.02	1.03	1.05

These data were each converted to a triangular probability density function (PDF), and then these were averaged to form the overall PDF. Finally, the overall PDF was integrated to construct the cumulative density function (CDF).

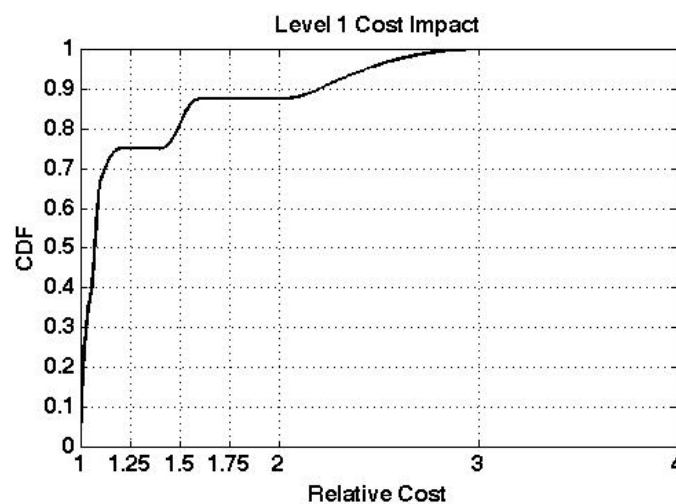



Figure 11.4-1. Estimate of Cost to Achieve Level 1 on the A2 Scale

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 61 of 112

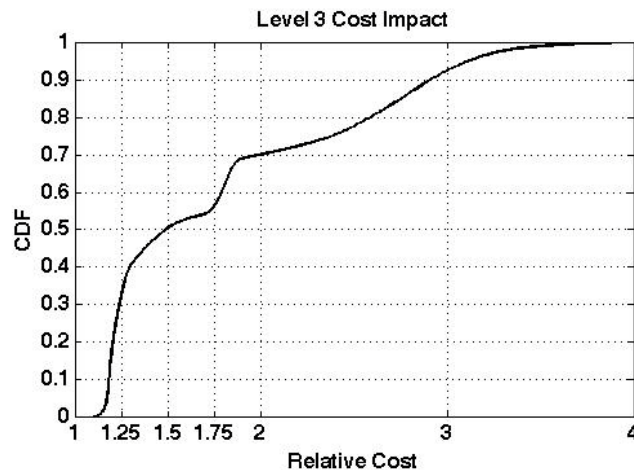



Figure 11.4-2. Estimate of Cost to Achieve Level 3 on the A2 Scale

[Figure 11.4-1](#) shows the CDF for the estimated cost to achieve Level 1, and [Figure 11.4-2](#) does the same for achieving Level 3. Although these estimates are not for the credibility assessment scale in the Permanent M&S Standard, the estimates for Level 1 on the A2 scale do apply directly to achieving Level 0 on the final credibility assessment scale. The reason is that Level 1 on the A2 scale and Level 0 on the final credibility assessment scale correspond to merely satisfying the requirements in the M&S Standard, which themselves only require documentation and reporting. The estimate for Level 3 on the A2 scale is roughly comparable to achievement of Level 3 on the credibility assessment scale, because both require substantial activities over and above the documentation and reporting requirements.

The cost estimates for merely satisfying the documentation and reporting requirements, shown in [Figure 11.4-1](#), suggest that in two-thirds of the cases, the additional cost to the M&S project will be less than 10 percent. Furthermore, the M&S projects that reported the least cost impact turned out to be those that are used in major development projects, whereas those that report the most cost impact are closer to the research code stage. The cost impacts shown in [Figure 11.4-2](#) are much more substantial. But, presumably such a high level on the credibility assessment scale would only be required for the most critical decisions.

These cost estimates are only theoretical, but they are based on all the data that was submitted during the pilot studies. Information on the actual cost impact of the M&S Standard awaits practical experience with it.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 62 of 112


12.0 Lessons Learned

Lessons learned from Phase 2 include:

1. Developing an M&S Standard that covers all types of models and simulations and all phases of the modeling and simulation process is extremely challenging.
2. Of the various challenges of developing a scale, i.e., picking an architecture, choosing the factors, and writing the level definitions, the hardest is writing clear, objective level definitions. This is easily overlooked by those new to such an activity.
3. Trained facilitation was extremely useful in containing the passionate “discussions” about the scale.
4. Once a decision is made, the temptation to revisit that decision is only contained by a firm rule requiring a formal motion accompanied by a second to even to begin the discussion.
5. Pilot studies are very important in bringing practical experience to bear on the development of a new standard.
6. The supermajority rule for final decisions is critical to ensuring that the final product had consensus support from the Topic Working Group.
7. Dedicated funding (as opposed to a volunteer activity) and involvement of practitioners was extremely beneficial to ensuring a feasible standard that would be accepted by the M&S community.
8. A high-level champion, in this case the OCE, was indispensable to overcoming barriers.

13.0 Definition of Terms


Corrective Actions	Changes to design processes, work instructions, workmanship practices, training, inspections, tests, procedures, specifications, drawings, tools, equipment, facilities, resources, or material that result in preventing, minimizing, or limiting the potential for recurrence of a problem.
Finding	A conclusion based on facts established by the investigating authority.
Lessons Learned	Knowledge or understanding gained by experience. The experience may be positive, as in a successful test or mission, or negative, as in a mishap or failure. A lesson must be significant in that it has real or assumed impact on operations; valid in that it is factually and technically correct; and applicable in that it identifies a specific design, process, or decision that reduces or limits the potential for failures and mishaps, or reinforces a positive result.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 63 of 112

Observation	A factor, event, or circumstance identified during the assessment that did not contribute to the problem, but if left uncorrected has the potential to cause a mishap, injury, or increase the severity should a mishap occur. Alternatively, an observation could be a positive acknowledgement of a Center/Program/Project/Organization's operational structure, tools, and/or support provided.
Problem	The subject of the independent technical assessment/inspection.
Proximate Cause	The event(s) that occurred, including any condition(s) that existed immediately before the undesired outcome, directly resulted in its occurrence and, if eliminated or modified, would have prevented the undesired outcome.
Recommendation	An action identified by the assessment team to correct a root cause or deficiency identified during the investigation. The recommendations may be used by the responsible Center/Program/Project/Organization in the preparation of a corrective action plan.
Root Cause	One of multiple factors (events, conditions, or organizational factors) that contributed to or created the proximate cause and subsequent undesired outcome and, if eliminated or modified, would have prevented the undesired outcome. Typically, multiple root causes contribute to an undesired outcome.

14.0 Acronyms List


AIAA	American Institute of Aeronautics and Astronautics
ARC	Ames Research Center
ASME	American Society of Mechanical Engineers
CAIB	Columbia Accident Investigation Board
CDF	Cumulative Density Function
COTS	Commerical-Off-The-Shelf
DFRC	Dryden Flight Research Center
DMSO	Defense Modeling and Simulation Office
DoD	Department of Defense
DoE	Department of Energy
EDL	Entry, Descent and Landing
EMB	Engineering Management Board
FY	Fiscal Year
GOTS	Government-Off-The-Shelf
GRC	Glenn Research Center
GSFC	Goddard Space Flight Center

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 64 of 112


ICD	Interface Configuration Document
IM&S	Integrated Modeling and Simulation
JPL	Jet Propulsion Laboratory
LaRC	Langley Research Center
M&S	Models and Simulations
MER	Mars Exploration Rover
MOTS	Modified-Off-The-Shelf
MSFC	Marshall Space Flight Center
NASA	NASA Aeronautics and Space Administration
NESC	NASA Engineering and Safety Center
NPD	NASA Policy Document
NPR	NASA Procedural Requirement
NRB	NESC Review Board
OCE	Office of the Chief Engineer
PCMM	Predictive Capability Maturity Model
PDF	Probability Density Function
RCC	Reinforced Carbon-Carbon
RPG	Recommended Practices Guide
RTF	Return To Flight
SRL	Simulation Readiness Level
SSC	Stennis Space Center
TSPO	Technical Standards Program Office
TSWG	Technical Standards Working Group
VPMM	Validation Process Maturity Model
VV&A	Verification, Validation and Accreditation

15.0 References


1. A Renewed Commitment to Excellence: An Assessment of the NASA Agency-wide Applicability of the Columbia Accident Investigation Board Report. B2005-100968, January 30, 2004 [URL: http://www.nasa.gov/pdf/55691main_Diaz_020204.pdf]
2. Return to Flight Task Group, Return to Flight Task Group Final Report, U. S. Government Printing Office, August 2005 [URL: http://returntoflight.org/assets/pdf/final_rfttg_report/RTF_TG_Final_Report_Lo-Res.pdf, July 2005]
3. Columbia Accident Investigation Board Report. Vol. 1, August 2003 [URL: <http://caib.nasa.gov/>]
4. Standard for Models and Simulations, NASA-STD-(I)-7009, Dec. 1, 2006
5. Standard for Models and Simulations, NASA-STD-7009, July 11, 2008

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 65 of 112

6. *VV&A Recommended Practices Guide*, Defense Modeling and Simulation Office.
<http://vva.dmsso.mil/>. RPG Build 3.0
7. *Guide for Verification and Validation of Computational Fluid Dynamics Simulation*. AIAA G-077, AIAA: Reston, VA, 1998
8. *Guide for Verification and Validation in Computational Solid Mechanics*. ASME V&V 10, ASME. New York, NY, 2006
9. Pilch, M.; Trucano, T.G.; Peercy, D.E.; Hodges, A.L.; Froehlich, G.K.: "Concepts for Stockpile Computing." SAND2004-2479, Sandia National Laboratories, 2004 (Official Use Only document, not available electronically)
10. Fasanella, E.L.; Jackson, K.E.; Lyle, K.H.; Jones, L.E.; Hardy, R.C.; Spellman, R.L.; Carney, K.S.; Melis, M.E.; Stockwell, A.E.: "Dynamic Impact Tolerance of Shuttle RCC Leading Edge Panels Using LS-DYNA." AIAA-2005-3631, 2005
11. Green, L.L.; Blattnig, S.R.; Luckring, J.M.; Tripathi, R.: "An Uncertainty Structure Matrix for Models and Simulations." AIAA-2008-2154, 2008
12. Harmon, S.Y.; Youngblood, S.M.: A proposed model for simulation validation process maturity. **J. Defense Modeling & Simulation**, vol. 2, no. 4, pp. 179–190, 2005
13. Oberkampf, W.L.; Pilch, M.; Trucano, T.G.: "Predictive Capability Maturity Model for Computational Modeling and Simulation." SAND2007-5948, Sandia National Laboratories, 2007
14. Lin, J.; West, J.S.; Williams, R.W.; Tucker, P.K.: "CFD Code Validation of Wall Heat Fluxes for a GO₂/GH₂ Single Element Combustor." AIAA-2005-4524, 2005
15. Blattnig, S.R.; Green, L.L.; Luckring, J.M.; Morrison, J.H.; Tripathi, R.K.; Zang, T.A.: "Towards a Credibility Assessment of Models and Simulations." AIAA-2008-2156, 2008
16. Hale, J.P.; Hartway, B.L.; Thomas, D.A.: "A Common M&S Credibility Criteria-set Supports Multiple Problem Domains." The 5th Joint Army-Navy-NASA-Air Force (JANNAF) Modeling and Simulation Subcommittee Meeting, CDJSC 49, May 2007, Johns Hopkins University: Columbia, MD
17. Mehta, U.B.: Guide to credible computer simulations of fluid flows. **J. Prop & Power**, vol. 12, no. 5, pp. 940–948, 1996.
18. Fogg, B.J.; Tseng, H.: "The Elements of Computer Credibility." Proceedings of CHI 99, pp. 80–87, ACM: New York, 1999
19. Tseng, S.; Fogg, B. J.: Credibility and computing technology. **Comm. ACM**, vol. 42, no. 5, pp. 39–44, 1999


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 66 of 112

20. Balci, O.: “Quality Assessment, Verification, and Validation of Modeling and Simulation Applications.” Proceedings of the 2004 Winter Simulation Conference. ed. by Ingalls, R.G.; Rossetti, M.D.; Smith, J.S.; Peters, B.A., pp. 122–129, Dec. 5-8, 2004, IEEE: Piscataway, NJ
21. Wang, R.Y. and Strong, D.M.: Beyond accuracy: what data quality means to data consumers. **J. Mgmt. Info. Systems**, vol. 12, no. 4, pp. 5–34, 1966.
22. Mehta, U.B.: “Simulation Credibility Level.” The 5th Joint Army-Navy-NASA-Air Force (JANNAF) Modeling and Simulation Subcommittee Meeting, CDJSC 49, May 2007, Columbia, MD: Johns Hopkins University.
23. Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information.” **Psychological Rev.**, vol. 63, pp. 81–97, 1956.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 67 of 112

Volume II: Appendices

- Appendix A. NASA Chief Engineer Memo
- Appendix B. Excerpts from the Space Shuttle Return to Flight Report
- Appendix C. ARC Topic Working Group Member's Objections to the Interim M&S Standard
- Appendix D. Consequence Definitions from NPR 8000.4
- Appendix E. Traceability Matrix
- Appendix F. Pilot Scale Questionnaire
- Appendix G. Decision Maker Interview Guide
- Appendix H. Pilot Questionnaire

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 68 of 112

Appendix A. NASA Chief Engineer Memo

September 1, 2006

Office of the Chief Engineer

TO: Distribution

FROM: Chief Engineer


SUBJECT: NASA Standard for Models and Simulations (M&S)

The NASA Engineering and Safety Center (NESC) is coordinating the development of a NASA Standard which will provide criteria for use of M&S. Addressing the findings and recommendations from the Columbia Accident Investigation Board, we need requirements that will improve our ability to develop, validate, and maintain computer models. The Chief Engineer has an action to establish Agency M&S requirements by the end of fiscal year 2006.

The M&S Standard will

- Ensure that the credibility of M&S results is properly conveyed to those making critical decisions,
- Assure that the credibility of M&S meet the project requirements,
- Establish M&S requirements and recommendations that will form a strong foundation for disciplined (structure, management, control) development, validation and use of M&S within NASA and its contractor community,
- Include a standard method to assess the credibility of the M&S presented to the decision maker when making critical decisions (i.e., decisions that effect human safety or mission success) using results from M&S,
- Establish a common set of terms and a uniform way for M&S practitioners to communicate the credibility of M&S,
- Be responsive to Diaz Action #4.

The standard development is in its final stages, with only a few, but critical, issues left to be resolved. As issuance of this standard is time critical—it is my desire that it be completed before the end of September 2006. I am requesting that the Topic Working Group members from the Centers make this their top priority for the rest of this fiscal year. In practice this means that they participate fully (via telecon) in all decisional meetings. I anticipate that there will be weekly

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 69 of 112

decisional meetings during September. I ask the Engineering Management Board members to ensure that their Centers' Topic Working Group members can make this commitment, and, if not, to identify an alternate or replacement.

Questions related to the processing of the draft M&S Standard should be addressed to Tom Zang, Thomas.A.Zang@nasa.gov at the NASA Langley Research Center.

Original signed by


Christopher J. Scolese

Enclosure

Distribution:

Engineering Management Board (EMB) Members

Ames Research Center/Doty, Laura
Dryden Flight Research Center/Stoliker, Patrick C.
Glenn Research Center/Gonzalez-Sanabria, Olga D.
Goddard Space Flight Center/Figueroa, Orlando
Jet Propulsion Laboratory/Muirhead, Brian
Johnson Space Center/Altemus, Stephen J.
Johnson Space Center/Watkins, Bobby J.
Kennedy Space Center/Wiley, Warren I.
Kennedy Space Center/Simpkins, Pat
Langley Research Center/Sandford, Stephen P.
Marshall Space Flight Center/Rudolphi, Michael U.
Stennis Space Center/Hebert, Bartt J.
Stennis Space Center/Rodriguez, Miguel A.
HQ/Scolese, Christopher J.
HQ/Robinson, Gregory L

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 70 of 112

HQ/Roe, Ralph R.

HQ/Lyles, Garry

HQ/Fishkind, Stanley

HQ/Ledbetter, Kenneth W.

HQ/Oconnor, Bryan

HQ/Ross, Harriet

HQ/Weinstein, Richard

HQ/Sorrels, Carrie

Ames Research Center/Unmeel B. Mehta

Glenn Research Center/Jeffrey (Jeff) J. Rusick

Goddard Space Flight Center/Thomas (Tom) V. McCarthy

Johnson Space Center/Galen P. Overstreet

Kennedy Space Center/Martin J. Steele

Langley Research Center/Steve R. Blattnig

Langley Research Center/Richard (Dick) E. Davis

Langley Research Center/Lawrence L. Green

Langley Research Center/James M. Luckring

Langley Research Center/Joseph (Joe) H. Morrison

Langley Research Center/Ram K. Tripathi


Langley Research Center/Thomas (Tom) Zang

Marshall Space Flight Center/Maninderpal (Paul) Gill

Marshall Space Flight Center/Joe Hale

Stennis Space Center/Jody Woods


HQ/Richard (Dick) Weinstein

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 71 of 112

Appendix B. Excerpts from the Space Shuttle Return to Flight Report

In July 2005, the Space Shuttle RTF Task Group issued their report [ref. 2]. Annex A2 contained numerous concerns about the use of M&S. The following are some excerpts from that report (bold face not in original but used here to highlight important points):

- “Standard engineering practice calls for objectives (**requirements and interface definitions**) to be established prior to development for any model or system of models, and processes and **criteria defined for validating and verifying the model's results**. ... Initially, we did not observe these normal processes being followed during the development of these models ...”
- “The uncertainties in one model (or system) inherently feeds into and compounds the uncertainty in the second model (or system), and so on. It appears, however, that NASA largely designed these five classes of models without the attention to the interdependencies between the models necessary for a complete understanding of the end-to-end result. Understanding the characteristics of, and validating and verifying, one type of model without examining the implications for the end-to-end result is not sufficient. ... But, as the Columbia accident showed, in a high risk environment that involves many unknowns like human space flight, **experience and instinct are poor substitutes for careful analysis of uncertainty**.”
- “... during the return-to-flight effort, there has been an enormous expenditure of time and resources - amounting to tens of millions of dollars - without the discipline of a formal **development plan**, clear objectives, **explicit plans for verification and validation**, thorough **outside review**, documented ICDs between models, or a good understanding of the limitations of analytical systems employing **multiple, linked deterministic models**. Validation and verification planning has been left to the end of the process rather than the beginning. ... Analytical models have essentially driven the return-to-flight effort; however, industry and academic **standards and methods for developing, verifying, and validating the models have not been used**. In addition, no sensitivity analyses had been conducted and no empirical data from flight history had been incorporated in the models or their validation”

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 72 of 112

Appendix C. ARC Topic Working Group Member's Objections to the Interim M&S Standard

Minority Report

Unmeel B. Mehta

October 11, 2006

Summary


1. The degree of accuracy of simulation results is not assessed. The stated goals and Chris Scolese's priority requirement - "Clearly credibility of results is the requirement" (08/31/06) – require assessment of accuracy of simulation results. The credibility of simulation capability is assessed. There should be two scales—one for credibility of simulation results and the other for credibility of simulation capability.
2. The conversion from scale A2 or A3 to the summary scale is not judicious. The definitions of levels in Section 4.7.1 do not properly represent the result of assessment in Appendix A.
3. Scales A2 and A3 have different assessment items and different definitions for levels. Only one scale must be used in Appendix A for consistency.
4. Requirement 4.8.3-c is inappropriate. The exercise of this requirement significantly diminishes the worth of this standard because the stated goals are not addressed.

Supporting Information

Section 4.7

1. Chris Scolese wants to know the bottom line – the credibility of simulation results – "Clearly credibility of results is the requirement" (Aug. 31, 2006). He requested that a scale be developed to provide the level of credibility of simulation results – the rigor scale. The stated goals of the standard are for the credibility of simulation results. Section 4.7 does not address the degree of accuracy of the M&S result, but it addresses the degree to which the accuracy of the M&S result is known. The latter leads to the credibility of M&S capability. The reproducibility of results and the repeatability of the process are also addressed. Hence, this section deals with "Assessing the Credibility of M&S Capability." Again, Chris needs to know what is the credibility of simulation results presented to him - that is, the degree of accuracy of those results.
2. Both A.2.3 and A.3.4 require reporting of the lowest score achieved as the summary assessment (Requirement 4.7.1). That could lead to an inappropriate conclusion for simulation result credibility. For example, if "Process Control" and other categories in Figure A.2.2-1 were, respectively, Level 1 (red) and Level 4 (green), then the summary result would be labeled as "1" (red).


Appendix A.1

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 73 of 112

1. The number of categories and definitions of levels appreciably differ in scales A2 and A3. Only one multi-dimensional scale should be provided to assess the simulation capability.
2. The definitions for levels used in scales A2 and A3 are not the same, and these definitions do not correspond to those in Section 4.7.1. A number from 1 to 4 is reported to section 4.7.1 from A2 (Appendix A.2.3) or A3 (Appendix A.3.4), and this number assumes the meaning provided for that number in Section 4.7.1. The judiciousness of this mapping is questionable. For example, A3 mentions “peer review” at Level 3 for 14 of 15 categories and “external audit” at Level 4 for 8 of 15 categories. Other levels in A3 do not use the word “review.” Section 4.7.1 labels Level 4 as “endorsed.” In this section, at each level the word “reviewed” is used. This example exhibits a serious inconsistency in mapping from multi-dimensional assessment to a summary one-dimensional assessment. Likewise, “working, in progress,” the Level 1 definition for 14 categories of A3 (or “ad hoc” the definition of Level 1 for 6 categories of A2) is not same as “Research” (Level 1) defined in Section 4.7.1. If Level 1 is achieved in all categories of A2 or A3, then the summary definition of Level 1 cannot be as defined.
3. The definition of each level contains the word “uncertainty” (Section 4.7.1). However, Requirement 4.8.3 allows for the option of stating that no quantitative or qualitative value of uncertainty is available. The exercise of this option makes the summary credibility scale significantly less useful, and drastically diminishes the worth of this standard for achieving the stated goals. Note that AIAA Editorial Policy Statement on Numerical and Experimental Accuracy (January 1994) states that “the AIAA journals will not accept for publication any manuscript reporting numerical solutions of an engineering problem that fails to adequately address the accuracy of the computed results or experimental results unless the accuracy of the data is adequately presented.” NASA’s programs (costing millions to billions of dollars) must mandate reporting of simulation uncertainties and test uncertainties for all simulation-based and test-based decisions affecting safety and/or mission success. The option offered in Requirement 4.8.3-c is unacceptable.

Appendix A.2

1. “Adequacy of the M&S results for the desired application depends on requirements and detailed knowledge of the specific application; adequacy is not addressed in this section.” This statement also confirms the assertion that Section 4.7 and A2 assess the M&S capability, but not the credibility of simulation results for intended uses. Adequacy is relevant. For example, validation cannot be conducted, if numerical uncertainties are comparable or larger than uncertainties in experimental data. Chris needs to know the credibility of simulation results presented to him. Additionally, the word “validation” is defined with the phrase “from the perspective of the intended uses of the model” (page 11). A2 eliminates this phrase. Specific applications or intended uses are not addressed.
2. Just because Level 4, as defined, is achieved for solution verification, does that mean that simulation results are credible for the intended uses? Similar questions are asked for validation,


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 74 of 112

predictive capability, and technical review. Again, just because a formal external review is conducted does not necessarily lead to the conclusion that simulation results are credible for the intended uses. The outcome of this review determines the credibility of simulation result. As defined, Levels 1 to 4 indicate increase in credibility of M&S capability. Processes are assessed.

Appendix A.3

1. A3 has three basic items (a) Is it built for the intended use (IU)? (b) Is it well built (BW)? (c) Is it used right (UR)? Item (a) provides M&S capability assessment for the intended uses. Item (b) determines process maturity by addressing the quality of the construction of the M&S. Item (c) assesses how well M&S was used for intended uses, including operator/analyst proficiency. Essentially, simulation capability, including simulation processes and use of capability, is assessed. How is the credibility (accuracy) of simulation results determined from these assessments? Instead of the considered three axes, only one axis labeled “Are Uncertainties Quantified?” or “Are simulation results credible?” is relevant to assess credibility of simulation results for intended uses.

2. The linkage between the three dimensions IU, BW, UR and the levels in the Summary Credibility Scale is provided. However, the equivalence of the definitions of levels in A3 and those in the Summary Scale or the justification that the definitions of levels in the Summary Scale correctly represent the result of assessment with scale A3 is missing.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 75 of 112

Appendix D. Consequence Definitions from NPR 8000.4


NPR 8000.4 was undergoing revision as the task was completed. Since the revised version of this NPR may not have the same (or any) risk matrix that was used as the basis of the Consequence definitions which appear in Appendix A of the M&S Standard (discussed here in Section 7.6.2), we record here the verbatim text of Section 2.3.1.1 of the version of NPR 8000.4 that had an effective date of April 25, 2002.

2.3.1.1 Consequence.

Consequence is an assessment of the worst credible potential result(s) of a risk. The measurement units differ depending on the specific risk. For example, the consequence of a cost risk may correspond to specific dollar amounts or percentages of the program/project budget or the consequence of schedule risks may correspond to the length of time delays. Consequence classifications are defined generally as Catastrophic, Critical, Marginal, and Negligible. A sample classification approach might be as follows

- a. Class I - Catastrophic. A condition that may cause death or permanently disabling injury, facility destruction on the ground, or loss of crew, major systems, or vehicle during the mission; schedule slippage causing launch window to be missed; cost overrun greater than 50 percent of planned cost.*
- b. Class II - Critical. A condition that may cause severe injury or occupational illness, or major property damage to facilities, systems, equipment, or flight hardware; schedule slippage causing launch date to be missed; cost overrun between 15 percent and not exceeding 50 percent of planned cost).*
- c. Class III - Moderate. A condition that may cause minor injury or occupational illness, or minor property damage to facilities, systems, equipment, or flight hardware; internal schedule slip that does not impact launch date; cost overrun between 2 percent and not exceeding 15 percent of planned cost.*
- d. Class IV - Negligible. A condition that could cause the need for minor first aid treatment but would not adversely affect personal safety or health; damage to facilities, equipment, or flight hardware more than normal wear and tear level; internal schedule slip that does not impact internal development milestones; cost overrun less than 2 percent of planned cost.*

Note: The portions of these classifications concerning safety are defined within NPR 8715.3, "NASA Safety Manual."


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 76 of 112

Appendix E. Traceability Matrix


The first matrix below provides a mapping from the Agency-level decisions that initiated development of this standard to the specific requirements in the M&S Standard. The driving objectives were taken from **Diaz Action #4** plus the direction from the NASA Chief Engineer to include the credibility assessment scale.

For the six objectives in columns 2–7, which were taken from Diaz Action #4, green fill color indicates that that M&S Standard requirement has a strong correlation with the driving objective, and yellow fill color indicates a modest correlation. A green color appears in the last column only if the entire requirement is driven solely by the credibility assessment scale, and a yellow color indicates that a portion of the requirement is driven solely by the credibility assessment scale. In the latter case, the portion of the requirement driven by the credibility assessment scale is highlighted in yellow in the first column.


The traceability of a half-dozen recommendations is also provided in the second matrix. These particular recommendations tie directly to one or more of the explicit objectives of Diaz Action #4. However, these were deemed not suitable for requirements because they were not always achievable.

	NASA Engineering and Safety Center Technical Report	Document RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 77 of 117


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.1.1 – Shall document the risk assessment for any M&S used in critical decisions.	Y	Y	G	Y	Y	Y	
Req. 4.1.2 – Shall identify and document those M&S that are in scope.	Y	Y	G	Y	Y	Y	

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 78 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
<p>Req. 4.1.3 – Shall define the objectives and requirements for M&S products including the following</p> <ul style="list-style-type: none"> a. The acceptance criteria for M&S products, including any endorsement for the M&S. b. The rationale for the weights used for the subfactors in the credibility assessment scale (see Appendix B.4). c. Intended use. d. Metrics (programmatic and technical). e. Verification, validation, and uncertainty quantification (see section 4.4). f. Reporting of M&S information for critical decisions (see section 4.8). g. CM (artifacts, timeframe, processes) of M&S. 		G	G	G			Y

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 79 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.1.4 – Shall develop a plan (including identifying the responsible organization(s)) for the acquisition, development, operation, maintenance, and/or retirement of the M&S.		Y			G		
Req. 4.1.5 – Shall document any technical reviews performed in the areas of Verification, Validation, Input Pedigree, Results Uncertainty, and Results Robustness (see Appendix B).							G
Req. 4.1.6 – Shall document M&S waiver processes.		G	Y				

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 80 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.1.7 – Shall document the extent to which an M&S effort exhibits the characteristics of work product management, process definition, process measurement, process control, process change, and continuous improvement, including CM and M&S support and maintenance.			Y		G		
Req. 4.2.1 – Shall document the assumptions and abstractions underlying the conceptual model, including their rationales.	Y	G	Y				
Req. 4.2.2 – Shall document the basic structure and mathematics of the model (e.g., reality modeled, equations solved, behaviors modeled, conceptual models).		G	Y				

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 81 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.2.3 – Shall document data sets and any supporting software used in model development and input preparation.			G		Y		
Req. 4.2.4 – Shall document required units and vector coordinate frames (where applicable) for all input/output variables in the M&S.			G		Y		
Req. 4.2.5 – Shall document the limits of operation of models.	G		Y				
Req. 4.2.6 – Shall document any methods of uncertainty quantification and the uncertainty in any data used to develop the model or incorporated into the model.		G	Y				
Req. 4.2.7 – Shall document guidance on proper use of the model.	G		Y		Y		

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 82 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.2.8 – Shall document any parameter calibrations and the domain of calibration.		G	Y				
Req. 4.2.9 – Shall document updates of the model (e.g., solution adjustment, change of parameters, calibration, and test cases) and assign unique version identifier, description, and the justification for the update		G	Y		Y		
Req. 4.2.10 – Shall document obsolescence criteria and obsolescence date of the model.			Y		G		
Req. 4.2.11 – Shall provide a feedback mechanism for users to report unusual results to model developers or maintainers.						G	

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 83 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.2.12 – Shall maintain (conceptual, mathematical and computational) models and associated documentation in a controlled CM system.			Y		G		
Req. 4.2.13 – Shall maintain the data sets and supporting software referenced in Req. 4.2.3 and the associated documentation in a controlled CM system.			Y		G		

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 84 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
<p>Req. 4.3.1 – Shall do either of the following</p> <p>a. Ensure that simulations are conducted within the limits of operation of the models, or</p> <p>b. Placard the simulation and analysis results with a warning that the simulation may have been conducted outside the limits of operation and include the type of limit that may have been exceeded, the extent that the limit might have been exceeded, and an assessment of the consequences of this action on the M&S results.</p>	G						

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 85 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.3.2 – Shall document and explain any observed warning and error messages resulting from the execution of the computational model.			G				
Req. 4.3.3 – Shall document which computational models were used (including revision numbers) in the simulation.			G				
Req. 4.3.4 – Shall document the versions of M&S results.			G				
Req. 4.3.5 – Shall document data used as input to the simulation, including its pedigree (see Appendix B).			Y		G		Y
Req. 4.3.6 – Shall document any unique computational requirements (e.g., support software, main memory, disk capacities, processor, compilation options).			Y		G		

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 86 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.3.7 – Shall document the processes for conducting simulations and analyses for generating results reported to decision makers.		Y	G				
Req. 4.3.8 – Shall document the use history of M&S in the same or similar applications, which are relevant for establishing the credibility of the current M&S application (see Appendix B).							G
Req. 4.3.9 – Shall document the assessment as to the appropriateness of the simulation and analysis relative to its intended use.			G				
Req. 4.3.10 – Shall document the rationale for the setup and execution of the simulation and analysis.			G				

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 87 of 112

Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.4.1 – Shall document any verification techniques used and any domain of verification (e.g., the conditions under which verification was conducted).		G	Y				
Req. 4.4.2 – Shall document any numerical error estimates (e.g., numerical approximations, insufficient discretization, insufficient iterative convergence, finite-precision arithmetic) for the results of the computational model.		G	Y				
Req. 4.4.3 – Shall document the verification status of (conceptual, mathematical and computational) models.		G	Y				

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 88 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.4.4 – Shall document any techniques used to validate the M&S for its intended use, including the experimental design and analysis, and the domain of validation.		G	Y		Y		
Req. 4.4.5 – Shall document any validation metrics and referents, and data sets used for model validation.		G	Y		Y		
Req. 4.4.6 – Shall document any studies conducted and results of model validation.		G	Y		Y		

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 89 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
<p>Req. 4.4.7 – Shall document any uncertainty quantification processes used for the following</p> <ul style="list-style-type: none"> a. The referent data. b. The input data. c. The M&S results. d. The propagation of uncertainties. e. The quantities derived from M&S results 		G	Y				

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 90 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.4.8 – Shall document any quantified uncertainties, both physical and numerical, including the following a. The referent data. b. The input data. c. The M&S results. d. The propagation of uncertainties. e. The quantities derived from M&S results.		G	Y				
Req. 4.4.9 – Shall document the extent and results of any sensitivity analyses performed with the M&S.		G	Y				
Req. 4.5.1 – Shall identify and document the Recommended Practices that apply to M&S for the program/project.	G		Y		Y		

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 91 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.6.1 – Shall determine the depth of required training for developers, operators, and analysts				G			
Req. 4.6.2 – Shall document a. Training topics required for developers, operators, and analysts of M&S. b. Process and criteria for verifying that training requirements are met.			Y	G			
Req. 4.6.3 – Shall determine the qualifications for developers, operators, and analysts.				G			
Req. 4.7.1 – Shall assess the credibility of M&S results for each of the eight factors in the credibility assessment scale described in Appendices B.2 and B.3.							G

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 92 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.7.2 – Shall justify and document the credibility assessment for each of the eight factors referenced in Req. 4.7.1.							G
Req. 4.7.3 – Shall perform the roll-up to an overall score according to the process described in Appendix B.4.							G

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 93 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
<p>Req. 4.8.1 – Reports to decision makers shall include explicit warnings for any of the following occurrences, accompanied by at least a qualitative estimate of the impact of the occurrence</p> <ul style="list-style-type: none"> a. Any unachieved acceptance criteria (as specified in Req. 4.1.3 (a)). b. Violation of any assumptions of any model (as specified in Req. 4.2.1). c. Violation of the limits of operation (as specified in Req. 4.2.5). d. Execution warning and error messages (see Req. 4.3.2). e. Unfavorable outcomes from the intended use and setup/execution assessments (described in Req. 4.3.9 and Req. 4.3.10). f. Waivers to any of the requirements in this standard. 	<div style="background-color: #00FF00; text-align: center; font-size: 2em; font-weight: bold;">G</div>						

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 94 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
<p>Req. 4.8.2 – Reports to decision makers of M&S results shall include an estimate of their uncertainty and a description of any processes used to obtain this estimate as defined in Req. 4.4.7 and Req. 4.4.8.</p> <p>a. Reported uncertainty estimate shall include one of the following</p> <p>(1) A quantitative estimate of the uncertainty in the M&S results, or</p> <p>(2) A qualitative estimate of the uncertainty in the M&S results, or</p> <p>(3) A clear statement that no quantitative or qualitative estimate of uncertainty is available.</p>		G					

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 95 of 112


Requirements	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Req. 4.8.3 - Reports to decision makers shall include the level of credibility for M&S results and the subfactor weights, using the process specified in section 4.7.							G

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 96 of 112


Critical Recommendations	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
Rec. 4.3b: CM records should contain test cases that span the limits of operation for the M&S defined by the program or project. "Test cases" are defined as benchmark input/output sets used to verify proper execution of the M&S.	G			Y			
Rec. 4.3c: The simulation should fail in a manner that prevents misuse and misleading results. (1) The simulation should provide messages that detail the failure mode and point of failure. (2) The analyst should document and explain all failure modes, points of failure, and messages indicating such failures.	G		G		Y		

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 97 of 112

Critical Recommendations	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic	method to assess the credibility of the M&S
Rec 4.5j: Recommended Practices for the following should be identified: Identify best practices for user interface design to constrain the operation of the simulation to within its limits of operations.	G						
Rec. 4.6a: Recommended training topics for developers, operators, and analysts of M&S include: The intended use of limits of operation for models.	G				Y		
Rec. 4.6d: Recommended training topics for developers, operators, and analysts of M&S include: How to recognize unrealistic results from simulations.						G	
Rec. 4.6e: Recommended training topics for developers, operators, and analysts of M&S include: Feedback processes to				Y		G	

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 98 of 112

Critical Recommendations	knowledge of operations is captured in the user interfaces	tool verification and validation, certification,	documentation, configuration management, and quality assurance	training or certification requirements	tool management, maintenance, and obsolescence	user feedback when results appear unrealistic n	method to assess the credibility of the M&S
improve M&S processes and results, including providing feedback for results that are not credible, are unrealistic, or defy explanation.							

	NASA Engineering and Safety Center Technical Report	Document RP-08-118	Version: 1.0
Title:	M&S Standard Completion		Page #: 99 of 117

Appendix F. Pilot Scale Questionnaire

Background

This survey is intended to collect feedback on the Simulation Credibility Scale described in Section 4.7 and Appendix A of the [Interim] NASA Standard for Models and Simulations. The primary goal of the M&S Standard is to ensure that the credibility of the results from models and simulations is properly conveyed to those making critical decisions based in part on the M&S results.

For example, suppose that the setting is a Flight Readiness Review, and that the decision-makers are confronted with conflicting results for different simulations. For example, suppose that both Organization A and Organization B have performed analyses using M&S for this flight, and that the results from the M&S of Organization A predict a 1 in 20 chance of failure, whereas the results from the M&S of Organization B predict a 1 in 200 change of failure. The decision-maker is faced with deciding which M&S result to believe, i.e., which result is more “credible.” The purpose of the Simulation Credibility Scale is to provide an objective means to make this assessment.

Developing such a scale is not simple. The Interim M&S Standard for Models and Simulations represents the initial thinking on this subject. The Topic Working Group for this Standard is keenly interested in feedback from M&S practitioners and decision-makers in order to develop a better scale for the Permanent M&S Standard.


The philosophy behind the scale in the Interim M&S Standard is expressed as follows at the start of Section 4.7: *“Credibility assessments of M&S results address, in order of importance, the degree to which the accuracy of a result is known, the reproducibility of the results, and the repeatability of the process. In this section, credibility includes the degree to which the accuracy of the M&S result is known and not the degree of accuracy of the M&S result. Any measure of the achieved degree of accuracy of M&S results as compared to the required degree of accuracy is highly dependent on the specifics of the problem.”*

In Questions 17 and 23, you have the opportunity to critique this philosophy. Please respond to the remaining questions (1-16 and 17-22) in the context of the above philosophy.

Instructions

Please provide some background information on your M&S project and then provide the specific feedback requested below. At a later time, an additional survey will be taken that solicits feedback on other aspects of the Interim M&S Standard.

The term “**level**” refers to the “Credibility Levels” as used in Appendix A. The term “**category**” refers to the “credibility categories” as used in Appendix A2 and “credibility criteria” as used in Appendix A3.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 100 of 112

Your M&S project may consist of just a single model or of multiple models that are coupled or linked together. Part of the feedback concerns how well the scales work for results from an individual model, and part with how well the scales work for coupled models.

Please respond to the multiple-choice questions with the answer that most closely matches your opinion.

M&S Project Information

Name of M&S Program/Project

Type of M&S Use (see Tables 1 and 2)

- a. Operations
- b. Manufacturing, Assembly, Test, and Evaluation
- c. Design and Analysis
- d. Natural Phenomena Prediction
- e. Technology Investment
- f. Scientific Data Analysis
- g. Scientific Understanding
- h. Training and/or Education
- i. M&S Research
- j. other (describe)

Number of Workyears in Development

Number of Individual Models

Software Life-Cycle Model


- a. waterfall (http://en.wikipedia.org/wiki/Waterfall_model)
- b. iterative (http://en.wikipedia.org/wiki/Iterative_and_incremental_development)
- c. spiral (http://en.wikipedia.org/wiki/Spiral_model)
- d. agile (http://en.wikipedia.org/wiki/Agile_software_development)
- e. other (describe)

Responder

Center:


Name:

Role on M&S Project:


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 101 of 112

Questions

1. How understandable is the Scale in Appendix A2?
 - a. I can understand the full Scale well enough to apply it simply by reading the Interim M&S Standard. No coaching is required.
 - b. I can understand most but not all of the Scale just from reading the Interim M&S Standard. However, in order for me to apply it, a modest amount of coaching is required.
 - c. I cannot understand the Scale well enough to apply it without considerable coaching.
 - d. I cannot understand the Scale, even with a substantial amount of coaching.
2. What, if any, parts of Scale A2 are not understandable or vague?
3. How understandable is the Scale in Appendix A3?
 - a. I can understand the full Scale well enough to apply it simply by reading the Interim M&S Standard. No coaching is required.
 - b. I can understand most but not all of the Scale just from reading the Interim M&S Standard. However, in order for me to apply it, a modest amount of coaching is required.
 - c. I cannot understand the Scale well enough to apply it without considerable coaching.
 - d. I cannot understand the Scale, even with a substantial amount of coaching.
4. What, if any, parts of Scale A3 are not understandable or vague?
5. How easy is Scale A2 to score for a typical individual model in your M&S?
 - a. The Level Definitions are very clear: I can readily assign a unique level in each category.
 - b. The Level Definitions are mostly clear: I can assign a unique level in most categories, but in a minority of the categories, I am uncertain about which of 2 adjacent levels to choose.
 - c. The Level Definitions are vague: In most categories I am unclear which level to choose.
 - d. The Level Definitions are meaningless: I have virtually no idea which level to choose in any category.
6. How would you improve the Level Definitions for Scale A2? If you can improve, please provide your definitions of Levels.
7. How easy is Scale A3 to score for a typical individual model in your M&S?
 - a. The Level Definitions are very clear: I can readily assign a unique level in each category.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 102 of 112

- b. The Level Definitions are mostly clear: I can assign a unique level in most categories, but in a minority of the categories, I am uncertain about which of 2 adjacent levels to choose.
 - c. The Level Definitions are vague: In most categories I am unclear which level to choose.
 - d. The Level Definitions are meaningless: I have virtually no idea which level to choose in any category.
8. How would you improve the Level Definitions for Scale A3? If you can improve, please provide your definitions of Levels.
9. How well does Scale A2 work for coupled M&S? (By “coupled M&S” we mean simulations that link/integrate multiple individual models.)
 - a. The Scale applies equally well in all categories to coupled models as to single models.
 - b. The Scale applies equally well in the majority, but not all, of the categories to coupled models as to single models.
 - c. Major changes are needed for the Scale to work for coupled M&S
 - d. The Scale cannot possibly work for coupled M&S
10. How well does Scale A3 work for coupled M&S? (By “coupled M&S” we mean simulations that link/integrate multiple individual models.)
 - a. The Scale applies equally well in all categories to coupled models as to single models.
 - b. The Scale applies equally well in the majority, but not all, of the categories to coupled models as to single models.
 - c. Major changes are needed for the Scale to work for coupled M&S
 - d. The Scale cannot possibly work for coupled M&S
11. Does the result of applying Summary Credibility Scale in Section 4.7 based on the result of the scale in Appendix A2 provide the decision-maker with a useful measure of credibility of the results of M&S?
 - a. The result provides an excellent measure of credibility.
 - b. The result provides a good measure of credibility.
 - c. The result provides a poor measure of credibility, but can be modified to provides a good measure of credibility
 - d. The result provides a poor measure of credibility, and a completely different approach is required.


	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 103 of 112

12. Does the result of applying Summary Credibility Scale in Section 4.7 based on the result of the scale in Appendix A3 provide the decision-maker with a useful measure of credibility of the results of M&S?
 - a. The result provides an excellent measure of credibility.
 - b. The result provides a good measure of credibility.
 - c. The result provides a poor measure of credibility, but can be modified to provides a good measure of credibility
 - d. The result provides a poor measure of credibility, and a completely different approach is required.


13. Given the first paragraph of Section 4.7, the Summary Credibility Scale, and the recipe to go from scale A2 to the Summary Scale, how well is the primary goal of the M&S Standard satisfied? [The primary goal is to ensure that the credibility of the results from models and simulations is properly conveyed to those making critical decisions]
 - a. The Scale does an excellent job of satisfying the primary goal.
 - b. The Scale does a good job of satisfying the primary goal.
 - c. The Scale does a poor job of satisfying the primary goal, but can be modified to do a good job.
 - d. The Scale does a poor job of satisfying the primary goal, and a completely different approach is required.

14. Given the first paragraph of Section 4.7, the Summary Credibility Scale, and the recipe to go from scale A3 to the Summary Scale, how well is the primary goal of the M&S Standard satisfied? [The primary goal is to ensure that the credibility of the results from models and simulations is properly conveyed to those making critical decisions]
 - a. The Scale does an excellent job of satisfying the primary goal.
 - b. The Scale does a good job of satisfying the primary goal.
 - c. The Scale does a poor job of satisfying the primary goal, but can be modified to do a good job.
 - d. The Scale does a poor job of satisfying the primary goal, and a completely different approach is required.

15. How well does Summary Credibility Scale based on the scale in Appendix A2 satisfy the secondary goal of the M&S Standard? [The secondary goal is to assure that the credibility of the results from M&S meets the project requirements]
 - a. The Scale does an excellent job of satisfying the secondary goal.
 - b. The Scale does a good job of satisfying the secondary goal.
 - c. The Scale does a poor job of satisfying the secondary goal, but can be modified to do a good job.

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 104 of 112


- d. The Scale does a poor job of satisfying the secondary goal, and a completely different approach is required.
16. How well does the Summary Credibility Scale based on the scale in Appendix A3 satisfy the secondary goal of the M&S Standard? [The secondary goal is to assure that the credibility of the results from M&S meets the project requirements]
 - a. The Scale does an excellent job of satisfying the secondary goal.
 - b. The Scale does a good job of satisfying the secondary goal.
 - c. The Scale does a poor job of satisfying the secondary goal, but can be modified to do a good job.
 - d. The Scale does a poor job of satisfying the secondary goal, and a completely different approach is required.
 17. What are the categories that most contribute (in your mind) to the credibility of simulation results presented to the decision-maker? [Include in your response any related comments you have on the philosophy that is italicized on p. 1 of this questionnaire.]
 18. What is a good range for the number of distinct categories in a Scale? [The word “category” means “credibility category” as used in Appendix A2 and “credibility criterion” as used in Appendix A3.]
 - a. 1–2
 - b. 3–5
 - c. 6–10
 - d. 11–20
 19. The scale in Appendix A3 has a hierarchical structure, whereas the scale in Appendix A2 is non-hierarchical. What is your opinion about the use of a hierarchical Scale? (Ignore the details of the categories and level definitions in these particular scales.)
 - a. I prefer a hierarchical scale
 - b. I do not have an opinion
 - c. I prefer a non-hierarchical scale
 - d. I don’t under the distinction and need additional information
 20. The Summary Credibility Scale (Section 4.7) produces a single number. The scales in appendices A2 and A3 have multiple categories for each level. For producing a single number from these multiple categories do you favor
 - a. Giving all categories equal weight
 - b. I do not have an opinion
 - c. Weighting some categories more than others

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 105 of 112

21. The Summary Credibility Scale (Section 4.7) produces a single number. The scales in appendices A2 and A3 have multiple categories for each level. For producing a single number from these multiple categories do you favor
 - a. Choosing the minimum
 - b. Choosing the simple average
 - c. Choosing a weighted average
 - d. Reporting the mean, minimum and maximum over all the categories

22. Both the scales in appendices A2 and A3 use a color-coded scheme to report the comparison with the required level and the achieved level in each category. A green-yellow-red color coding is used to denote whether the achieved level is equal or greater than, exactly one level less than, or two or more levels less than the required level.
 - a. This color coding is useful
 - b. I do not have an opinion
 - c. This color coding is not useful

23. What other comments would you care to make on the subject of the Simulation Credibility Scale? [Include in your response any remaining comments you have on the philosophy that is italicized on p. 1 of this questionnaire.]

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 106 of 112

Appendix G. Decision Maker Interview Guide

M&S Credibility Interview Guide Introduction

Interviewee: _____ Date: _____

Title: _____

Telephone No: _____

Interviewers: _____

Introduction


We are conducting interviews to examine some major M&S activities—both successes and failures. The M&S activities of interest are those used to support program/project decisions that affect human safety and mission success. We want to understand in each case what worked, what did not work, and why. We are particularly interested in your thoughts on the major contributors to the credibility of simulation results (fourth item above).

We can use your input in exploring one of those cases. To this end, could you comment on the last major model/ simulation that you dealt with, with respect to the following areas? We intend for this to take only a small amount of your time, but the more detail you can afford us, the better the credibility communication system we can construct for your use.

[Note: This section also contained the Diaz Action #4 language and the list of objectives from the NASA Chief Engineer memo of Sept. 1, 2006.]

Interview Guide

1. Briefly describe the last important M&S model output you were asked to consider.
 - a. Case Name: (What was it? Commit to a particular instance.)
 - b. Please describe the overall project supported by the M&S activity. (Describe its context in general terms.)
 - c. Briefly describe the system, disciplines or phenomena that were modeled. (Describe the narrower context of the modeled system in general terms. Get enough information that you can infer the general class—see attached appendix for a list of classes.)
 - d. What were the particular “results” of the M&S that were used to support the decisions? (What were the outputs you used, in specific terms?)

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 107 of 112

e. Please describe the role of the M&S results in the decision. (How did you use the outputs in the decision?)

2. We now would like to focus on the credibility of the results.

a. Were the M&S results credible? If so, why? If not, why not? (Identify the key factors.)

b. How much time did you give to think about its credibility?

c. Did you have enough information to determine its credibility?

d. What additional information could you have used to determine its credibility?

e. Do you think it would have been possible to explore its credibility more fully, given the time and resources available?


f. *Please review this list of factors and identify those that made a major contribution to the credibility or lack thereof of the M&S results. [This question is optional at the discretion of the Center Topic Working Group member. If this question is asked, (a) be sure to clearly distinguish the answer to this question from the answer to #2a, and (b) use whatever list you deem appropriate.]*

g. Would a credibility scale attached to an M&S output be useful to you? (Closed question; ask following if response is “yes”)

h. What concerns do you have regarding a credibility scale?

3. Are there any other questions I should have asked?

4. Is there anyone else I should speak with?

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 108 of 112

Appendix H. Pilot Questionnaire

Instructions

This questionnaire seeks to obtain feedback on several key issues involving the M&S Standard. Answers to the questions in Sections A-C are required. Responses to the questions in Sections D are desirable but not required.

Please provide some background information on your M&S project and then provide the specific feedback requested below.

Please respond to the multiple-choice questions with the answer that most closely matches your opinion.

A. M&S Project Information

[Note: The same background information was collected as in the Pilot Scale Questionnaire (Appendix E), and the following instructions were added to identify the baseline for the cost estimates.]

M&S Project Baseline Cost

In Section B we are asking for your best estimate of the added cost to your M&S Project that would result from compliance with the M&S Standard. We'd like your cost estimate to be provided in percentage terms relative to your M&S Project Cost Baseline. Please provide this M&S Project Cost Baseline in terms of either the annual full cost or the life-cycle full cost of the M&S Project. Also provide a similar baseline for the entire program or project that your M&S project supports. Furthermore, we'd like to know whether your M&S project started with existing M&S or needed to develop new M&S capability as part of the development. Include both the development and operation phases of the M&S project in the cost estimate.

Annual Cost Baseline

Annual full cost of M&S development & operation


Annual full cost of the major program or project supported by the M&S project

Life-cycle Cost Baseline

Life-cycle full cost of M&S development & operation

Life-cycle full cost of the major program or project supported by the M&S project

New or existing M&S?

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 109 of 112

B. Achievement of Goals

The **primary goal** of this standard is to ensure that the credibility of the results from models and simulations (M&S) is properly conveyed to those making critical decisions. This will support risk-informed decisions. (By “critical decisions” we mean decisions that may affect human safety or project-defined mission success criteria.)


The **secondary goal** is to assure that the credibility of the results from M&S meets the project requirements. This will reduce the risks associated with critical decisions.

Since we have already collected your input on the Summary Credibility Scale (and the two supporting versions detailed in Appendices A2 and A3), please respond to the questions in this section without considering Section 4.7 and Appendix A.

1. How well is the primary goal of the M&S Standard satisfied?
 - a. The Standard does an excellent job of satisfying the primary goal.
 - b. The Standard does a good job of satisfying the primary goal.
 - c. The Standard does a poor job of satisfying the primary goal, but can be modified to do a good job.
 - d. The Standard does a poor job of satisfying the primary goal, and a completely different approach is required.
2. What aspects of the M&S Standard detract from achievement of the primary goal?
3. What aspects need to be added to the M&S Standard to achieve the primary goal?
4. How well is the secondary goal of the M&S Standard satisfied?
 - a. The Standard does an excellent job of satisfying the primary goal.
 - b. The Standard does a good job of satisfying the primary goal.
 - c. The Standard does a poor job of satisfying the primary goal, but can be modified to do a good job.
 - d. The Standard does a poor job of satisfying the primary goal, and a completely different approach is required.
5. What aspects of the M&S Standard detract from achievement of the secondary goal?
6. What aspects need to be added to the M&S Standard to achieve the secondary goal?

C. Cost-Benefit Analysis

The cost of complying with the M&S Standard can be strongly dependent upon the Simulation Credibility Level that is required by the program. Hence, we are asking the cost question with respect to achieving specified levels of credibility. To achieve Credibility Level 1 requires one merely to conform to the documentation and reporting requirements. This cost is independent of whether one uses Scale A2 or Scale A3 to determine Credibility. The cost of achieving

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 110 of 112

Credibility Levels 2–4 very likely does depend upon which Scale is used. In answering question 10 & 11, please pick either Scale A2 or Scale A3 as the basis for your answer.

A cost estimate such as this is very difficult to make. Therefore, we are asking you to supply 3 numbers for each cost estimate: the most likely cost, the minimum cost and the maximum cost.

Furthermore, state your cost numbers in percentage terms, using your M&S Project Cost Baseline, as documented above, as the baseline for each answer. In particular, do not provide the cost to achieve Level 3 Credibility as the incremental cost beyond achieving Level 1 Credibility. Instead, give the total cost to reach Level 3 Credibility compared to the M&S Project Cost Baseline.


7. Are your cost estimates based on Scale A2 or Scale A3?
8. What is the percentage additional cost (relative to the M&S Project Cost Baseline) imposed by complying with the Standard at Level 1 Credibility?
 - a. maximum percentage cost increase
 - b. most likely percentage cost increase
 - c. minimum percentage cost increase
9. What aspects of the Standard are the major cost drivers for reaching Level 1?
10. What is the percentage additional cost (relative to the M&S Project Cost Baseline) imposed by complying with the Standard at Level 3 Credibility?
 - a. maximum percentage cost increase
 - b. most likely percentage cost increase
 - c. minimum percentage cost increase
11. What aspects of the Scale are the major cost drivers for reaching Level 3?
12. Describe any benefits that you believe will occur because of compliance with the Standard (at Level 1 Credibility)

D. Optional Questions

Additional Cost Estimates

If you have the time to provide cost estimates for reaching Level 2 and Level 4 Credibility, please do so in Questions 13-16. Use the same choice of Scale as noted above in Question 7 for your answers.

13. What is the percentage additional cost (relative to the M&S Project Cost Baseline) imposed by complying with the Standard at Level 2 Credibility?

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 111 of 112

- a. maximum percentage cost increase
- b. most likely percentage cost increase
- c. minimum percentage cost increase

14. What aspects of the Scale are the major cost drivers for reaching Level 2?

15. What is the percentage additional cost (relative to the M&S Project Cost Baseline) imposed by complying with the Standard at Level 4 Credibility?

- a. maximum percentage cost increase
- b. most likely percentage cost increase
- c. minimum percentage cost increase

16. What aspects of the Scale are the major cost drivers for reaching Level 4?

Clarity

Since we have already collected your input on the Summary Credibility Scale (and the two supporting versions detailed in Appendices A2 and A3), please respond to the next 2 questions without considering Section 4.7 and Appendix A.

17. How understandable is the Standard?

- a. I can understand the Standard well enough to apply it simply by reading it. No coaching is required.
- b. I can understand most but not all of the Standard just from reading it. However, in order for me to apply it, a modest amount of coaching is required.
- c. I cannot understand the Standard well enough to apply to without considerable coaching.
- d. I cannot understand the Standard, even with a substantial amount of coaching.

18. What, if any, parts of the Standard are not understandable or vague?


General Comments

19. What other comments would you care to make on the subject of the M&S Standard?

Requirements Assessment

Ideally, each requirement in the Standard is

- Valid (necessary for the goals of the Standard)
- Verifiable (an independent part can determine objectively whether the requirement was met)
- Doable (achievable with sufficient training, time and money)
- Applicable (relevant to all types of M&S, e.g., empirical curve fits, partial differential equations, discrete event simulation, operations models)

	NASA Engineering and Safety Center Technical Report	Document #: RP-08-118	Version: 1.0
Title: M&S Standard Completion			Page #: 112 of 112

If you believe that any requirement in the M&S Standard fails to meet one of the 4 criteria list above, please provide the supporting details in the table below. You may also use the complete list of requirements provided in the additional attachment for your response.

Rqmt. Number	Valid?	Verifiable?	Doable?	Applicable?